UAH RI Report 2004-1

LESSONS FROM ARMY SYSTEM DEVELOPMENTS

VOLUME I
RESEARCH QUESTIONS AND ANALYSIS

William A. Lucas
The Cambridge-MIT Institute
Massachusetts Institute of Technology

Richard G. Rhoades
Research Institute
The University of Alabama in Huntsville

June 2004

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 2004	Final (September 29, 1999 – May 31, 2004)

4. TITLE AND SUBTITLE
LESSONS FROM ARMY SYSTEM DEVELOPMENTS: Volume I Research Questions and Analysis

5. FUNDING NUMBERS
C: DAAH01-98-D-R001 DO 67 and
DAAH01-97-D-R005 DO 49

6. AUTHORS William A. Lucas (MIT) and Richard G. Rhoades(UAH)

6. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research Institute
The University of Alabama In Huntsville
Von Braun Research Hall
Huntsville, AL 35899

The Cambridge-MIT Institute
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBER
UAH RI Report 2004-1

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Laboratory Demonstration Office
Aviation and Missile Research, Development and Engineering Center
Redstone Arsenal, AL 35898

DCG for Acquisition and Technology
US Army Materiel Command
9301 Chapek Road
Fort Belvoir, VA 22060-5527

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
None assigned

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
This report documents the results of a multi-year research project which employed a structured case study approach to examine the history and processes that had resulted in the introduction of a number of technology-based Army systems in time to make a positive contribution to the outcome of Desert Storm. In addition to the case studies documented in Volume II of this report, a common set of data was obtained and analyzed to identify factors contributing to successful systems development;(this) Volume I contains the results of this analysis.

Several of the statistically significant relationships found involve factors that are related to the stability of the program. When key members of the project team left the program too early, project outcome suffered. Further, both project funding cutbacks and project team turn-over negatively correlated with the quality of the testing program and the timeliness of key test events. These two attributes of the testing program also had the strongest correlation with project outcomes. In addition, changes in systems requirements during development correlated with poor project cost performance, and, finally, turn-over in key user representative personnel correlated negatively with system performance in the field. A central conclusion from this study is that shorter development cycle times favorably correlate with key project outcome variables, largely by minimizing the exposure of the project to destabilizing influences which have also been shown to correlate negatively with these same outcome variables.

14. SUBJECT TERMS acquisition strategy, case studies, weapon systems development, program stability, testing and simulation, technology maturity, project team structure, project leader attributes

15. NUMBER OF PAGES
70

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

This report documents the results of a research project of several years duration which employed a structured case study approach to examine the history and processes that had resulted in the introduction of a number of technology-based Army systems in time to make a positive contribution to the outcome of Desert Storm. Volume II of the report contains the 15 case studies that were developed on systems ranging from the M829A1 "silver bullet" to the GUARDRAIL Common Sensor and the APACHE attack helicopter.

The case studies were developed through the use of structured interviews with key participants from the government/contractor team that developed each system. In addition to the case studies, this process resulted in collection of a common set of data for the systems studied which could then be analyzed to identify factors contributing to successful system development. The results of this analysis are contained in (this) Volume I of the report. Two of the 15 case studies examined systems which might have been useful on the battlefield (based on the views of Army technical leaders), but that failed to successfully complete development. The intent of including failures in the research was to provide a basis for distinguishing factors which contributed to both successful and unsuccessful system developments. While they are useful for the qualitative lessons they offer, two cases are inadequate for quantitative analysis and most analysis focuses on the 13 successful cases. It is therefore an assessment of contributors to the relative degree of success.

All 15 systems studied are listed in Table 1.1, which follows on the next page. For each system, this table also contains information on the duration of the development phase of the program and a summary of the project manager's description of the most difficult problem encountered. It is interesting to note that lack of sustained user support for the requirement the system was intended to satisfy was mentioned as the most difficult problem for the two failures, but user-related issues were not identified for any of the successful development cases.

The heart of any systematic study is the definition of a common outcome measure that allows comparison. The obvious path was to compare the projects and systems based on their performance relative to their agreed upon goals and requirements. Each project had a budget, a systems procurement cost goal, a set of technical requirements, and completion dates. In addition, three questions of performance are immediately observable and easily remembered by project managers: Did the system go into production? Once production was started were problems found that required that further engineering changes be made? And did the system perform well in its use in Desert Storm? Structured questions were used to ask the key government and industry interviewees about how well their projects performed in these areas, with a range of answers that characterized how badly the projects had missed meeting their objectives, if they had not been completely successful.

Six of the outcome measures mentioned above were used to create a scale that scores the (system) projects from zero to six according to the number of key outcomes a project achieved. If a project was (1) transitioned to production on time, (2) developed within budget, (3) had no late engineering changes, met both (4) the goals for system unit costs and

(5) its technical requirements, and encountered (6) no difficulties when it was deployed in the field, it was awarded (the maximum) six points. These results also appear in Table 1.1.

System/case	Development duration (months)	PM's most difficult problem	Key outcomes achieved (0-6)
APACHE attack helicopter	108	Control of production costs; influenced by integration plant location choices	1
TADS/PNVS (target acquisition and designation/pilot's night vision systems)	~36	Cost growth in development	3
MLRS rocket system	33	Establishing and managing four nation cooperative development program	6
ATACMS missile system	37	Key vendor went out of business	6
M40 chemical protective mask	~48	Immaturity of critical technologies	2
Dismounted microclimate cooler **Note: Did not enter full development**	Not applicable	Lack of stable user requirements due to immaturity of technology	Not applicable
Mounted microclimate cooler	~24	Key vendor failed to support integration schedule	5
M829-A1 armor-piercing kinetic energy tank ammunition	~36	Achieving needed innovation in system design	6
FOG-M (fiber optic guided missile) **Note: Did not complete development**	Not applicable;	Lack of sustained user support	Not applicable
TOW-2A (Tube-launched missile)	48	Stability of threat armor requirements	3
AN/TAS 4 infrared night sight	~24	Selection of unqualified vendor and split management responsibility	4
Joint Stars Ground Station	105	Cost and schedule growth/delivering complex software	1
Guardrail common sensor	~24	Complexity of integration of mission equipment	3
PAC-2 (PATRIOT anti-missile system)	~52	Early fielding to meet SCUD missile threat	2
HELLFIRE missile system	~84	Adversarial relationship between key vendor and prime	3

Table 1.1 – Summary case information

Standard statistical analysis procedures appropriate for this number of cases and type of data were used to identify and evaluate correlations between the factors studied and the

several outcome variables, and, in some cases, among the factors. The results of these analyses are summarized in Table 1.2. The testing/simulation and technology maturity factors were included because of their identification in recent Government Accounting Office studies as key determinants of success.

Factor	Relationships Found/Comments
1. Project team characteristics and practices:	
--leadership	Team leader's perceived ability to obtain resources, his/her breadth of experience and ability to resolve technical issues all are positively related to reduced engineering changes during production and completing development within budget
--staffing	Low turnover in key project team members relates positively to completing development within budget, to meeting system unit cost targets and to achieving system performance objectives
2. Role of government S&T organizations	Army labs/centers were typically actively involved in both pre-development and development phases; actively involved in both successes and failures; and actively involved in both short and long developments
3. Testing and simulation approach	Validating component and system maturity at the right time in the program relates positively to completing development within budget, to meeting system unit cost targets and to successful performance in the field. The quality of the testing and simulation conducted relates positively to reduced engineering changes during production and to meeting system unit cost targets.
4. Importance of stability:	
--funding	Funding uncertainty was related to increased turnover in key project team members and the need to deal with changes in testing plans and other project structure issues
--system requirements	Changes in system requirements, particularly during the middle of development, relate to an increase in late engineering changes and negatively to project success in meeting its goals for systems costs.
--key user (TRADOC) personnel	Changes in key TRADOC personnel during development relates to less successful performance in the field
5. Timely communication of problems	Nearly all cases described timely communication of problems from contractor to government PM and from government PM to Army leadership.
6. Importance of technology maturity (TRLs)	Maturity of critical technologies used in systems studied, as measured by TRLs, was similar to that found in previous study of small electronics projects. No positive correlation found between higher TRLs at the start of development and most outcome variables

Table 1.2 - Summary of significant relationships

Several of the statistically significant relationships involve factors that are related to the stability of the program. When key members of the project team left the program too early, project outcome suffered. Further, both project funding cutbacks and project team turn-over negatively correlated with the quality of the testing program and the timeliness of key test events. These two attributes of the testing program also had the strongest correlation with project outcomes. In addition, changes in systems requirements during development

correlated with poor project cost performance, and, finally, turn-over in key user representative personnel correlated negatively with system performance in the field.

Taken together, these several relationships strongly suggest that stability of program resources and objectives is a very powerful influence on the relative success of the project. In reflecting on this array of instabilities that could impact a system development, it became clear that they had at least one thing in common: The longer a system stayed in development, the greater chance it had to experience one or more of these program destabilizing events. Or, stated another way, shorter system development cycles should result in better project outcomes. When this hypothesis was tested by examining the correlation between the system development durations and the aggregate outcome scale (See the data in Table 1.1), a strong correlation was found. A central conclusion from this study is therefore that shorter development cycle times favorably correlate with key project outcome variables, largely by minimizing the exposure of the project to destabilizing influences which have also been shown to correlate negatively with these same outcome variables.

Whether or not a change to selecting projects with shorter development times is made, the Army could do more to stabilize the guidance and resources given to both shorter and longer development projects. Acting alone, the Army could do more to map rotating personnel assignments and other sources of TRADOC change to project development cycles. It could eliminate all but the most critically important changes in systems requirements once projects move beyond early development since it appears that, as widely believed, such changes will almost certainly hurt project performance. Through contracts and informal management practices, the Army could work with its contractors to provide better continuity of development project staffing.

The defense acquisition community has long recognized that lengthy systems development times are disadvantageous. Sometimes the associated negatives have been phrased in program instability terms and this study certainly provides a strong empirical support for those who hold these beliefs. Over the years a number of initiatives have been attempted to shorten development cycles, with limited success where complex systems were involved. The current approach is referred to as "spiral development"; its basic concept is to get a useful, if limited, capability in the field quickly and then introduce additional technology-based capabilities through further "spirals" of development. This approach appears to be in keeping with the implications of this study's central conclusion.

2. INTRODUCTION

Background

Desert Storm was one of the most remarkable military conflicts ever fought. Its uniqueness is found in its one-sidedness: what could have been a protracted small war against an Iraqi force of 600,000 troops was concluded in 17 days of ground combat, with only 36 troops lost to enemy action. This was an historic triumph of training, organization, logistics and technology. In the specific case of the US Army, a number of new military systems, incorporating sophisticated technology, made their first significant battlefield appearance in Desert Storm.

This research project focuses on the process that brought that technology to the battle field in order to develop insights for planning and organizing for the continued generation of technology-based systems. In this first decade of the 21st Century it is evident that the system of defense laboratories, contractors and technology programs that produced Desert Storm's technology is being fundamentally changed. The end of the Cold War, the current focus on the Global War on Terrorism, and the perceived absence of other significant military threats to the security of the nation are, to some significant extent, resulting in the dismantling of the organization and process of U.S. defense technology development that produced the success of Desert Storm.

This work took advantage of a window of opportunity. Desert Storm is now distant enough to allow perspective, and to enable the use of widely known information about technologies that were previously classified. At the same time, its history is recent enough that the key players in the development of this technology are still available to provide their recollections and insights. New research can now examine the development of military systems used in Desert Storm to provide insight into the keys to success and failure at that time, capturing lessons that might inform the management of Army technology development in the future.

Research Methodology

As noted above, the basic intent of this research was to examine the history and processes that had resulted in the introduction of a number of technology-based Army systems in time to make a positive contribution to the outcome of Desert Storm. In order to be able to examine as many different systems as possible within the constraints of the funding available for the study, the authors proposed that a significant portion of the work would be performed using "free labor"; experienced defense personnel enrolled in military and academic institutions would execute the data collection portion of the research (as the subject for a thesis or research paper). Each was to use a consistent framework for collecting and presenting data; this framework, in the form of a "Case Study Checklist"-a research questionnaire, was prepared by the authors. This approach, referred to sometimes as a "structured thesis," has been used successfully at MIT for many years. It leaves the student important latitude to identify important issues not in the guiding structure, and the opportunity to reach independent conclusions while still contributing to a unified research structure. This construct

intended to benefit from the maturity and experience of senior students who were already familiar with defense processes and systems.

This planned student involvement approach was implemented with partial success in this project. Research for one-third of the cases was carried out by students who matched the a priori experience and background assumptions. Two of these students used their research on this project as the basis of Masters theses which they wrote during their graduate study at the Naval Postgraduate School, under collaborative arrangements with Postgraduate School faculty developed by the authors. Research on another third of the cases was carried out by graduate students at the University of Alabama in Huntsville who did not have previous knowledge of defense processes and systems. One of the authors attempted to compensate for this lack of background by providing a series of tutorial sessions on the defense acquisition process and organizational relationships during the course of their work. Also, one of these students researched three cases, over a two year period, and was able to use the acquisition process experience he gained in developing the first case to advantage on the latter two cases. The final third of the cases were researched by Professor Dan Sherman, of the University of Alabama in Huntsville faculty; Dr. Sherman was knowledgeable of Army acquisition processes and organizations from his prior research experience. Project resources originally earmarked to support collaboration with faculty at a larger number of educational institutions were reallocated to fund Dr. Sherman's involvement.

In short, it proved more difficult than anticipated to find Army military or civilian students enrolled in programs which required a research project, who could be interested, on a voluntary basis, in participating in this effort. As a result, all 15 cases were researched by individuals with ties of one sort or another to Huntsville, Alabama organizations, and (as will be discussed) their choice of systems to research resulted in somewhat greater coverage of missiles and aviation related systems.

Each individual researching a system (case) carried out interviews using the structured questionnaire with key participants from the government and contractor project management teams which had been responsible for developing, producing and fielding that system. The researcher was then responsible to synthesize two products, which he provided to the authors. The first product was an "integrated" questionnaire that documented his view of the most accurate answers to the questions, based on the more detailed interviews he had conducted, and giving appropriate weight to the interviewee best situated to know "truth" in a particular case. For example, in the event of disagreement in the individual responses to questions about the functioning of the contractor's design teams, researchers were instructed to give greater weight to the views of the contractor program manager. The results of analysis of these answers across the systems studied appear in this volume (Volume I) of the report.

The second product was a system case study, documenting in narrative form his insights on the key issues discussed during the interviews. At a minimum, he was asked to discuss the issues dealt with in the research questionnaire, but was encouraged to examine other issues in which he had particular interest, or which had been raised by the interview subjects. Development of this series of system case studies was intended to significantly increase the number available for use by defense acquisition students and educators. For several systems

(FOG-M, MLRS, PATRIOT), these new case studies explored issues that were substantially different from those contained in prior cases on the same systems, deepening the documentary coverage for that particular system. The system case studies appear in Volume II of the report.

Question	System	Technology	Questionnaire
Outcomes?	X	X	O1-O10
Production readiness?	X		Page 1, T3,H6, B4-B6, B8
Technology readiness?	X	X	Page 1, T5-T7
Importance of technology to prime?		X	Page 1; T4
Familiarity of prime with technology?		X	Page 1; T2,T3
Role of gov't S&T organization?	X	X	T8-T10, B11 Page 1
Role of S&T organization that developed technology?	X	X	Page 1 T8-T10
Timeline?	X		Page 1
Difficulties in integrating technology?		X	T3, H3, B1, B4-B8
User support? (or role of user?)	X		D18, F5-F6,W3-W5
Key Issue for PM?	X		I2
Timely problem disclosure?	X		D12, D16,D19
Requirements stability?	X		F7,W6,B13
Test approach used?	X	X	V1-V15
IPT approach used?	X		H2,H4-H5, D7, D9, D1, D13, D14, D16, D19, F4
Proper staffing of IPT?	X		H3, D3-D6,D8, D10
Design to manufacturing linkage?	X	X	F1-F3, F10-F13, W1-W2, W16-W18
Funding stability?	X		H1, D11, B2
Design to supplier linkage?	X	X	F20-F23, W26-W28, B10

Table 2.1 – Research questions examined

As was previously noted, use of a research questionnaire (a "Case Study Checklist") to guide the interviews was a critical aspect of the research methodology. This questionnaire was designed by the authors to provide coverage of a number of development process, organizational relationship, critical technology maturity and other issues that either the authors' prior experience or the management literature suggested might be relevant to

determining the relative success of projects. A portion of the questionnaire consisted of questions that were in common with a research instrument successfully used by one of the authors in a prior study of aerospace research projects. The draft questionnaire was tested by four former Army system project managers (whose former system responsibilities were not included in the systems chosen for this research project). Their responses provided valuable suggestions for clarifying the wording of a few of the questions, which was done in the final version, and they found that completing the questionnaire could be done in about 30 to 45 minutes. The final questionnaire is provided as an APPENDIX, and has been modified by inserting the responses to the questions. Table 2.1 contains a listing of research questions incorporated into the questionnaire. This list includes whether the question applies at the technology or system level because in addition to questions about the system as a whole, a set of questions focused on the component systems and technologies.

Systems Studied

As was earlier noted, the common feature of the system developments studied in this research is that each system first was employed in a significant way on the battlefield in Desert Storm. That, in turn, meant that for the most part development began on these systems during the 1980s. It was the intent that the systems studied include examples from the broad array of military systems for which the (original) research sponsor- The Army Materiel Command (AMC)-had responsibility. To achieve that intent, the following process was used to develop a list of candidate systems from which the researchers could select systems to study:

1. Each Director of an AMC Research, Development and Engineering Center was asked to nominate candidate systems from his commodity area (e.g. missiles, aviation, communications) that met the criterion of having first been successfully used in a significant way in Desert Storm. Each Director was also encouraged to discuss this question with project managers that his organization supported, and include their input. Each was further asked to nominate any systems which, in their judgment, would have been militarily useful in Desert Storm, but had failed to complete development. (Note: this process resulted in relatively few such failures being identified.)

2. The list of candidate systems that resulted was discussed with the AMC Deputy Commander (who was a veteran of Desert Storm) and his civilian Senior Executive Service deputy. Together they divided the approximately 40 candidate systems into two groups, reflecting priority for research attention. The systems studied in this project were taken from the first priority group.

3. As students were recruited to participate in developing case studies, they were initially allowed to choose systems on a "first come, first served" basis. Presumably because the students were affiliated with Huntsville, Alabama organizations, this approach resulted in essentially complete coverage of the missiles and aviation-related systems. In order to broaden the coverage, Dr. Sherman was requested to select one of the failure-to-complete-development systems and two systems that were neither missiles nor aviation-related. Because of the missile and aviation selections of the early participants, later participants were also encouraged to select systems that broadened the coverage of the AMC commodity line. Table 2.2 summarizes the systems that were selected for study in this research project.

System	Researcher	Commodity category
APACHE attack helicopter	Ference	Aviation
TADS/PNVS (target acquisition and designation/pilot's night vision systems)	Oelrich	Aviation
MLRS rocket system	Sherman	Missiles
ATACMS missile system	Romanczuk	Missiles
M40 chemical protective mask	Ruocco	Soldier support
Dismounted microclimate cooler **Note: Did not enter production**	Ruocco	Soldier support
Mounted microclimate cooler	Ruocco	Soldier support
M829-A1 armor–piercing kinetic energy tank ammunition	Mitchell	Ammunition
FOG-M (fiber optic guided missile) **Note: Did not enter production**	Sherman	Missiles
TOW-2A (Tube-launched missile)	Vessels	Missiles
AN/TAS 4 infrared night sight	Granone	Target acquisition
Joint Stars Ground Station	Sherman	Intelligence
Guardrail common sensor	Sherman	Intelligence
PAC-2 (PATRIOT anti-missile system)	Sherman	Missiles
HELLFIRE missile system	Johansen	Missiles

Table 2.2 – Systems studied

3. GENERAL FINDINGS

Systems characteristics

As noted in the INTRODUCTION, an effort was made to include in this study a range of technologies and systems that broadly represented the variety found in the Army Materiel Command-supported portfolio of systems. The list included better-known systems such as the Apache helicopter, the Patriot PAC2 defensive missile system, and the M829 A1 sabot tank round (the "silver bullet"). It also included systems that are largely unknown: The M40 gas mask, a personal vest cooling system, and GUARDRAIL Common Sensor. As was described earlier, care was taken to include systems that were never produced ("failures") such as the FOG-M; systems that encountered serious development problems, delays and cost over-runs; systems that once produced and deployed to the Iraqi theater were found to have operational problems; as well as uniformly successful systems. The goal was to provide a varied cross-section of Army systems developed in the 1980s that included a range of military functions, development processes and types of outcomes that would serve as an empirical base roughly representative of US Army systems of that era. For the reasons which have been noted, missile and aviation-related systems make up about half of the study sample.

Table 3.1 on the next page provides a list of the 15 projects, their development start dates, the maturity of the technology incorporated when development began using the Technology Readiness Level scale, and the duration of the development phase for each project. It might be noted that in some instances a system is built on an earlier system development attempt which failed to reach operational use, and start dates are difficult to define. When the respondents were unsure of the specific month in which development began, a mid-year start was assumed. This uncertainty is identified in Table 3.1 by use of an approximate (~) sign. Projects with a broad or multiple purposes sometimes had mixed success, as represented by the case included on micro-cooling vests which is treated as two projects: a successful project for a vehicle mounted system, and a failed project for dismounted use.

The development durations for the systems studied range from two years to nine years. Half the cases took three years or less and half required four years or more. (The median duration was 37 months.) It is interesting that those cases which had shorter development durations did not necessarily have higher levels of system technology maturity at the start of the development phase. For example, both the Joint Stars Ground Station and the GUARDRAIL Common Sensor had Technology Readiness Levels of 7 at the start of development, yet one took over four times as long to complete development. Note also that the three systems with the lowest Technology Readiness Levels all had relatively short development durations. Chapter 4 will provide a more quantitative discussion of the impact of technology maturity on project outcome variables.

System	Development start	TRL* at start of development	Development duration (months)
APACHE attack helicopter	1973	5	108
TADS/PNVS (target acquisition and designation/pilot's night vision systems)	1977	3	~36
MLRS rocket system	1977	4	33
ATACMS missile system	1986	4	37
M40 chemical protective mask	1983	7	~48
Dismounted microclimate cooler **Note: Did not enter production**	Not Applicable; Did not enter development.	Not applicable; development not completed.	Not applicable; development not completed.
Mounted microclimate cooler	1982	3	~24
M829-A1 armor–piercing kinetic energy tank ammunition	1985	5	~36
FOG-M (fiber optic guided missile) **Note: Did not enter production**	1988	7	Not applicable; cancelled prior to completing development
TOW-2A (Tube-launched missile)	1980	6	48
AN/TAS 4 infrared night sight	1979	3	~24
Joint Stars Ground Station	1984	7	105
Guardrail common sensor	1984	7	~24
PAC-2 (PATRIOT anti-missile system)	1986	6	~52
HELLFIRE missile system	1973	6	~84

*Note: TRL means Technology Readiness Level (scale 1 to 9); see APPENDIX

Table 3.1 – Selected characteristics

Role of Science and Technology Organizations

The system acquisitions which are the central topic of this research encompass a set of organizational relationships which, while different in detail, are consistent in broad terms. For each of the systems, there is typically a defense contractor responsible for the design, development and production of the system; and the government "customer" made up by three potentially relevant subgroups within the government customer community. For these Army systems these were:

a. An Army Project Office responsible for managing the contract between the government and the contractor (often identified by the contractor as the "customer")

b. The "user" typically represented by the Training and Doctrine Command (TRADOC) provides clarifying input on operator interface issues to support design and development. TRADOC is the advocate for the importance of the requirement that the system is intended to fulfill (which may be vital in keeping system funding adequate).

c. The S&T organization (typically an Army laboratory or research, development and engineering center). This organization:

   1. May have been the developer of technologies being used in the system development,

   2. May have provided personnel to the system Project Office specifically to help effect transfer of those technologies, and/or

   3. May be involved in simulation, testing or other activities supporting the system development.

Table 3.2 documents the extent of involvement of Army laboratories and research, development and engineering centers in the acquisition of the fifteen systems studied. In most of the cases an Army laboratory or research, development and engineering center had a leadership role in the technology and system concept development work carried out prior to the initial recognition of its potential as a system. PAC-2 and Joint Stars Ground Station are exceptions, with the prime contractor responsible for leadership during that phase. This prime contractor role for PAC-2 is consistent with the PATRIOT strategy of funding research at the prime contractor to support a continued series of block improvements as the threat (i.e. system requirements) changed. Similarly, the prime contractor and major supplier co-leadership roles identified for Joint Stars during this phase are consistent with the restructuring of the program to an Army/Air Force joint program that occurred.

During the period between initial recognition of system potential and start of system development, the laboratory/center role is more varied. In eleven cases a leadership role continued, which was usually shared with the system Project Office. In the remaining four cases the Army laboratory or center provided active support during this phase.

Once systems development started, the laboratory or center typically provided active support in requirements interpretation, system engineering, simulation or testing.

Note that the Army laboratory or center was as likely to be actively involved in those systems that did not make it to production as in those that did, or in those that had relatively short development durations, as those that took longer to complete development. Perhaps the most interesting thing which these data show is the extent to which the Army laboratories and centers are involved with the systems acquisition process; this is consistent with previous studies that have highlighted the criticality of the support to the systems acquisition process role for these organizations.

System	Pre-system concept involvement Army lab/center	Pre-development involvement Army lab/center	Army lab/center contributions to maturity of key technologies at start of development	Development involvement Army lab/center.
APACHE attack helicopter	lead/co-lead	lead/co-lead	2 of 3	active
TADS/PNVS (target acquisition and designation/pilot's night vision systems)	lead/co-lead	active	3 of 3	active
MLRS rocket system	lead/co-lead	active	3 of 3	active
ATACMS missile system	lead/co-lead	lead/co-lead (with DARPA)	2 of 3	less than active
M40 chemical protective mask	lead/co-lead	lead/co-lead	3 of 3	lead/co-lead
Dismounted microclimate cooler **Note: Did not enter production**	lead/co-lead	lead/co-lead	Not applicable (did not enter full development)	not applicable (did not enter full development)
Mounted microclimate cooler	lead/co-lead	lead/co-lead	3 of 3	lead/co-lead
M829-A1 armor–piercing kinetic energy tank ammunition	lead/co-lead	lead/co-lead	3 of 3	active
FOG-M (fiber optic guided missile) **Note: Did not enter production**	lead/co-lead	lead/co-lead	3 of 3	active
TOW-2A (Tube-launched missile)	lead/co-lead	lead/co-lead	3 of 3	active
AN/TAS 4 infrared night sight	lead/co-lead	lead/co-lead	1 of 2	active
Joint Stars Ground Station	active	active	1 of 3	active
Guardrail common sensor	lead/co-lead	lead/co-lead	0 of 3	lead/co-lead
PAC-2 (PATRIOT anti-missile system)	active	active	2 of 3	active
HELLFIRE missile system	lead/co-lead	lead/co-lead	1 of 3	active

Table 3.2 – Army lab/center involvement in systems acquisition

Project Manager's Most Difficult Problem

_During the development of the research questionnaire used in this project, discussions were held with staff members of the Defense Systems Management College (an element of the Defense Acquisition University). Following these discussions, question I2 - "What was the most difficult problem the Project Manager faced, how was the problem dealt with, and what was the impact of the problem on the project outcome?"- was added to the questionnaire. This question was intended to obtain responses on a topic of particular interest to the College, and by so doing make the case studies resulting from this project potentially of greater usefulness to the College for acquisition education purposes.

The replies obtained to this question in the interviews are summarized in Table 3.3. All 15 cases studied are included, including the two failure-to-reach-production cases (FOG-M and Dismounted Microclimate Cooler). It is interesting to note that lack of sustained user support for the requirement which the system was intended to satisfy was mentioned as the most difficult problem for the two failures, but user-related issues were not identified for any of the successful development cases.

Case	Most Difficult Problem	Solution/Impact
APACHE	Control of production costs/influenced by integration plant location choices	Use of Army and DOD "pressure" on contractor to influence decisions/minimized impact
TOW 2A	Stability of armor threat requirements	Flexible systems engineering process that accommodated changes/minimized impact
GUARDRAIL Common Sensor	Complexity of integration of mission equipment	Use of "integrated product team" approach/minimized impact
FOG-M	Lack of sustained user support	Program could not survive development cost growth
Joint Stars Ground Station	Cost and schedule growth/delivering complex software	Additional funding and time required
TADS/PNVS	Cost growth in development	Additional funding obtained
M40 Mask	Immaturity of critical technologies	Design modified to accommodate more mature technologies
M829A1 Round	Achieving needed innovation in the system design	Design iterations employed
PAC-2	Early fielding to meet SCUD missile threat	Rapid changes in software were made
Dismounted microclimate cooler	Lack of stable user requirements due to immaturity of technology	Development program not supported
Night Sight	Selection of unqualified vendor and split management responsibility	Vendor replaced and single PMO given full responsibility
Mounted microclimate cooler	Key vendor failed to support integration schedule	RDEC used to provide expedited integration of initial units
HELLFIRE	Adversarial relationship between key vendor and prime	Army PMO staff helped to facilitate needed communications/impact minimized
ATACMS	Key vendor went out of business	Replacement vendor selected and was intensively managed by on-site senior prime contractor manager
MLRS	Establishing and managing four nation cooperative development program	Significant involvement of program leadership minimized impact

Table 3.3 - Project Manager's "Most Difficult Problem"

Various types of problems with vendors were identified in four of the cases, while other difficult problems encountered ranged from cost growth and schedule delays to threat requirements instability to the complexity of defining and implementing a multi-national development program. It is clear from this set of data that successful development and production can occur in spite of the need to deal with delays in reaching production, development cost increases, complex management arrangements and the like.

Outcomes

The heart of any systematic study is the definition of a common outcome measure which allows comparison. Detailed information about costs and performance can be difficult to obtain and rarely leads to comparable measures, for example the problem of measuring a gas mask's performance with that of a missile. The obvious path was to compare the projects and systems based on their performance relative to their agreed upon goals and requirements. Each project had a budget, a systems procurement cost goal, a set of technical requirements, and completion dates. In addition, three questions of performance are immediately observable and easily remembered by project managers: Did the system go into production? Once production was started were problems found that required that further engineering changes be made? And did the system perform well in its use in Desert Storm? Structured questions were also used to ask the key government and industry interviewees about how well their projects performed in areas such as was the project completed on time, on budget and did it meet its technical requirements. The range of answers provided characterized how badly the projects had missed meeting their objectives, if they had not been completely successful.

This study design and the nature of these outcomes facilitated the analysis of "relative success," but can say little about complete failure. If a system did not meet the important objective of being put into production, most of the other outcomes either never occurred or cannot be judged. Difficulty in meeting technical requirements may have caused delays and cost over-runs, but if the project was cancelled, then it is hard to say what technical performance would have been, when it would have been completed and other outcomes, had it been continued. The fact that only two failed projects were included in the study means that there is a poor statistical base for comparing success and failure to reach production; insights on that question must rely on a qualitative reading of the case studies. Most of the analysis reported in this report uses the outcome data on the remaining 13 cases to isolate the factors that are related with how well the systems acquisition process went for projects that went into production and eventually the field.

The following several figures depict the results of these outcome assessments:

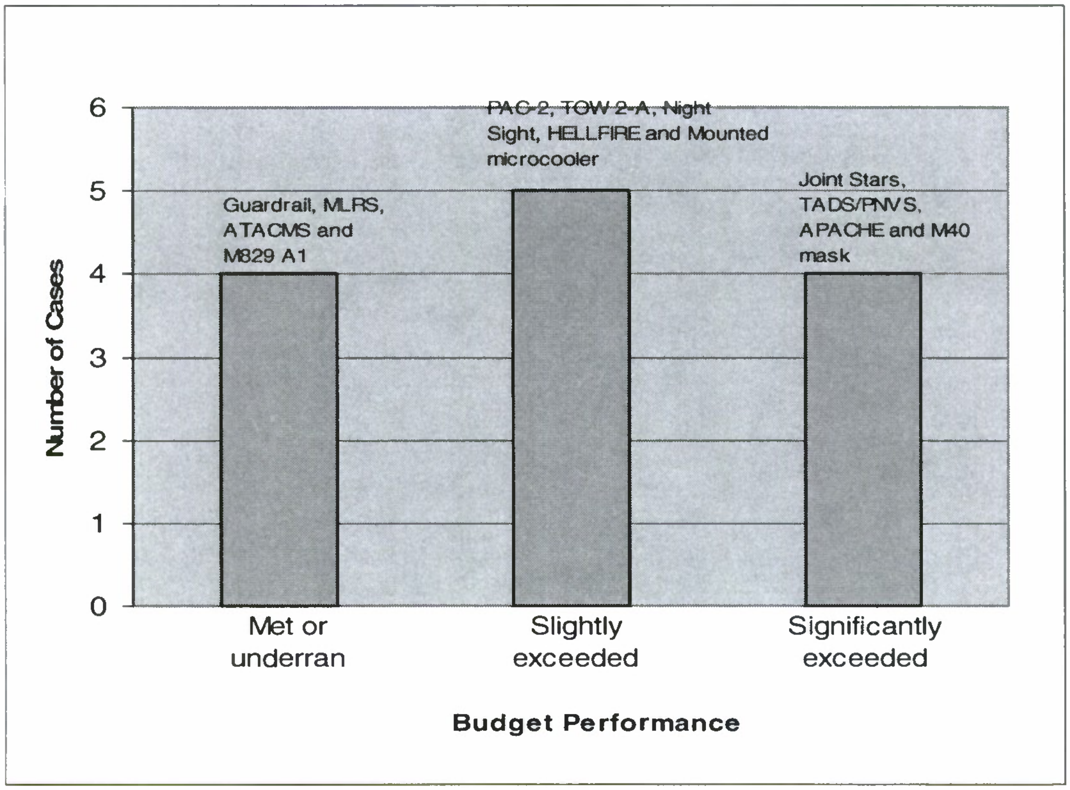

Figure 3. 1 – System development cost outcomes (O8)

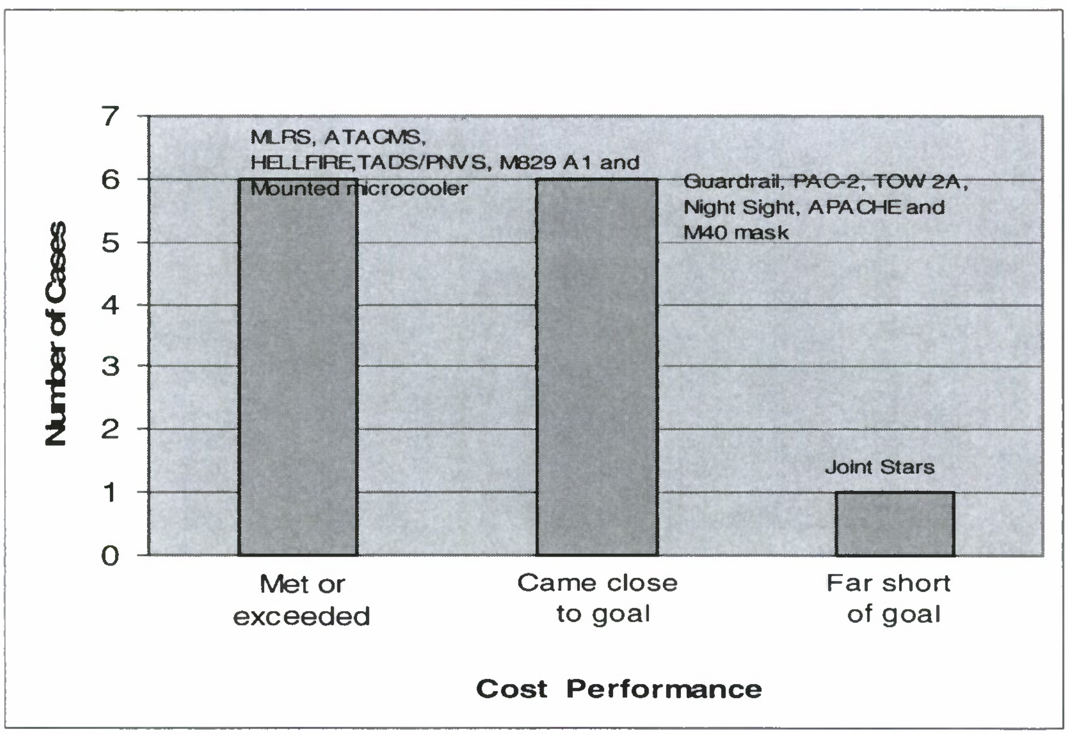

Figure 3. 2 – System unit cost outcome (O7)

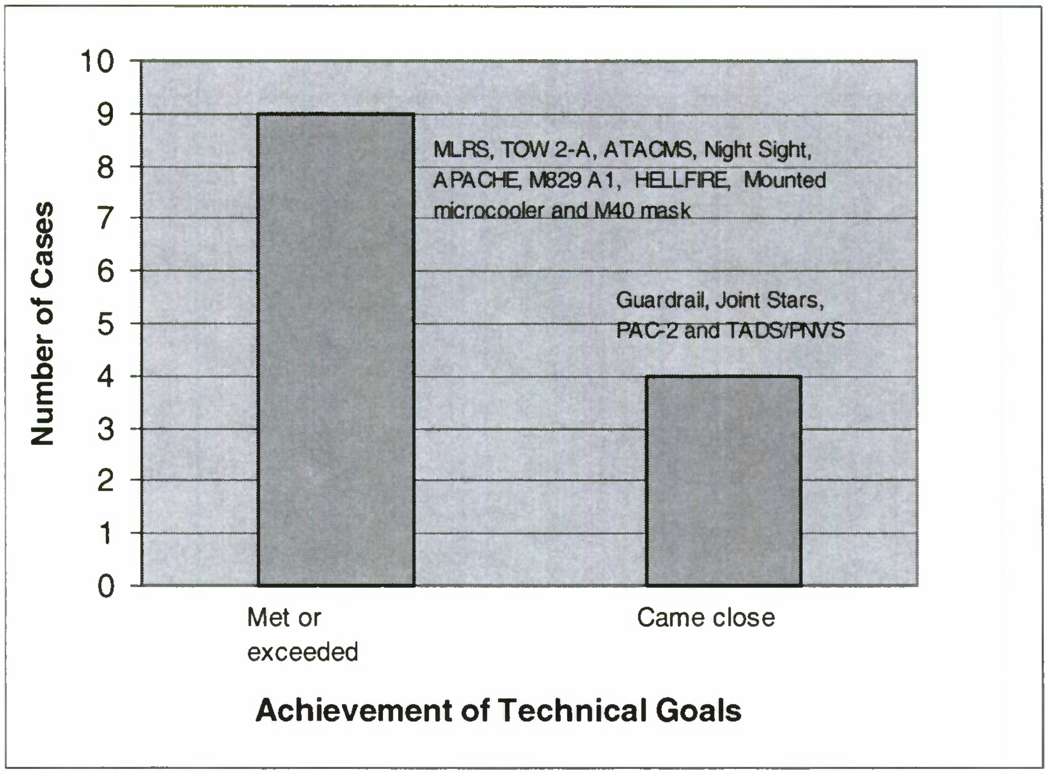

Figure 3.3 – System technical performance outcome (O9)

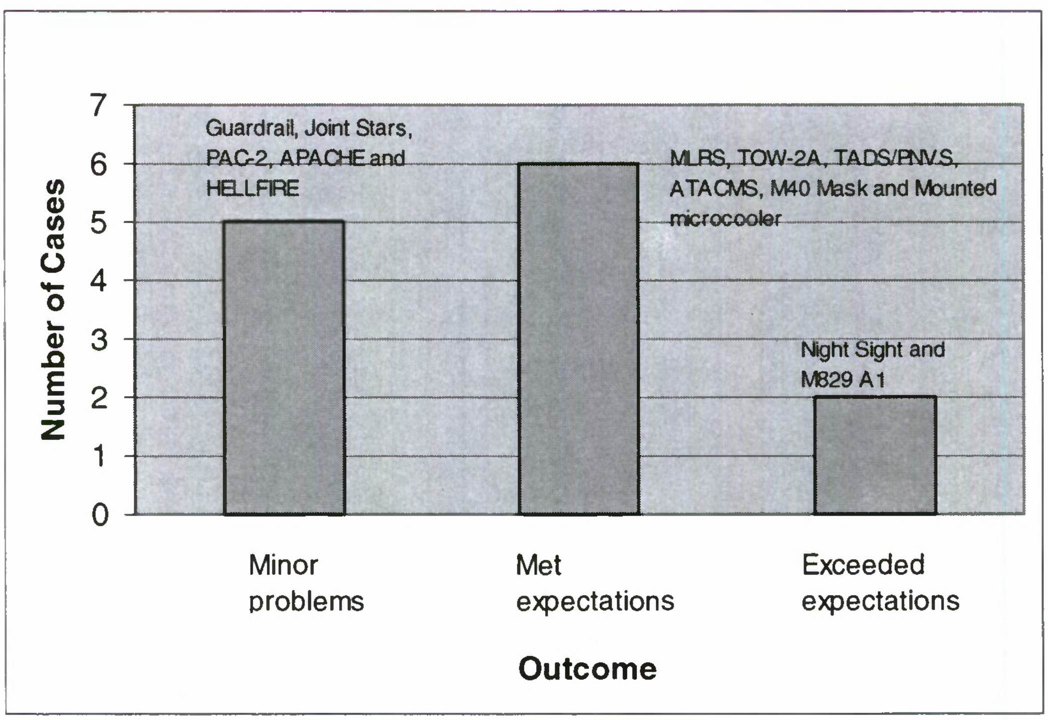

Figure 3.4 – Operational performance (O10)

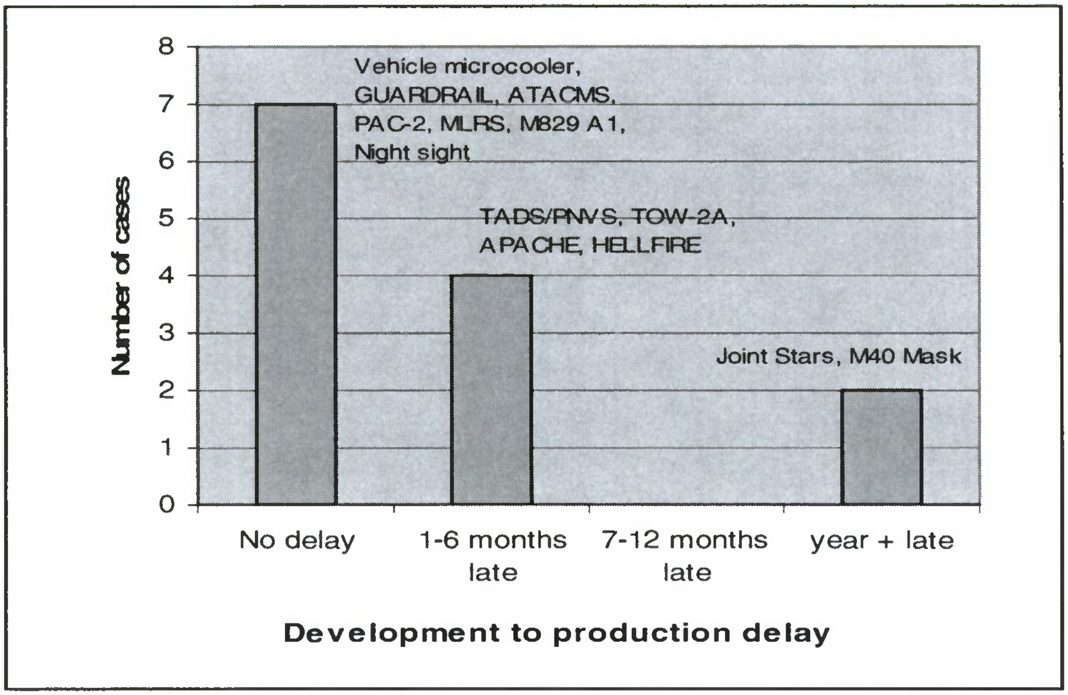

Figure 3.5 – Delay in transitioning to production (O5)

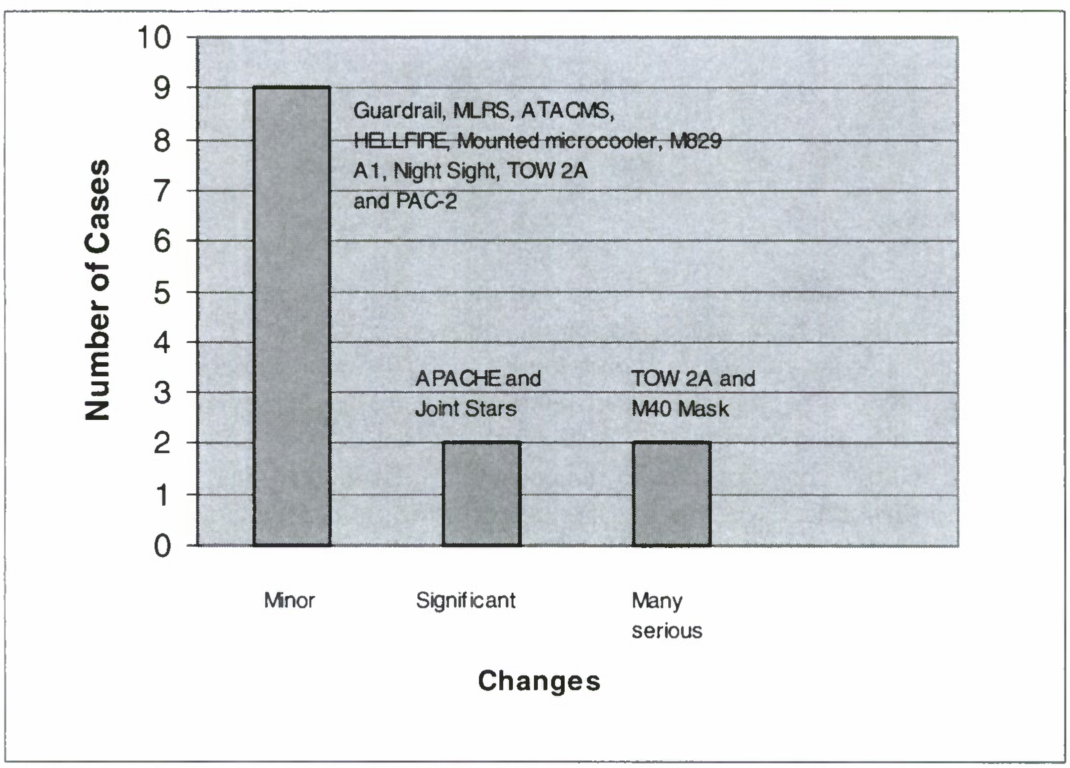

Figure 3.6 – Density of changes transitioning to production (O2)

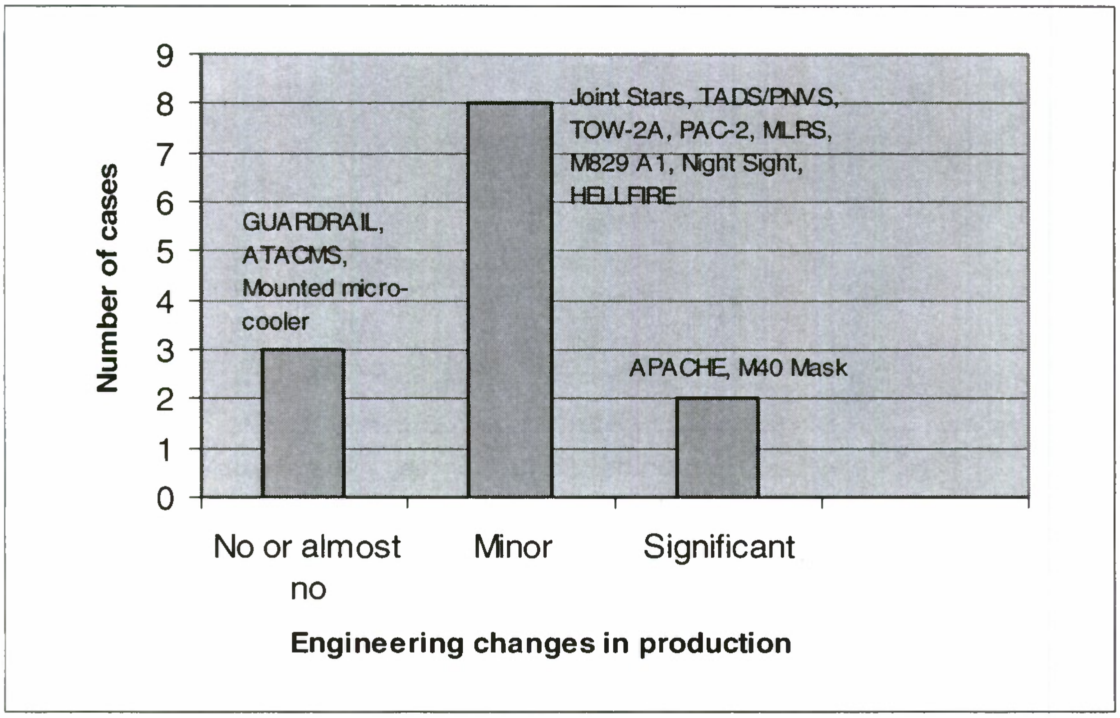

Figure 3.7 – Density of changes during production (O6)

Outcome scale

Many factors studied related to one or another of these outcomes. However, in an approach to assess their comparative importance, six of the outcomes previously shown (O5 – O10) were used to create a aggregate scale which ranks the (system) projects from zero to six according to the number of high performance outcomes a project achieved. If a project was (1) transitioned to production on time, (2) developed on budget, (3) had no late engineering changes, met both (4) the goals for system unit costs and (5) its technical requirements, and encountered (6) no difficulties when it was deployed in the field, it was awarded six points. Three of the projects were successful on all of these criteria and were given a score of six, while two projects achieved only one of these project goals and were each given a score of one. The mean score was 3.5. While the differences in outcome results presented here were rarely used as a basis for the statistical analysis which went into this report, comparing the means between groups is often used to provide a sense of the results in the body of the report. Figure 3.8, on the following page, depicts this scale in histogram form

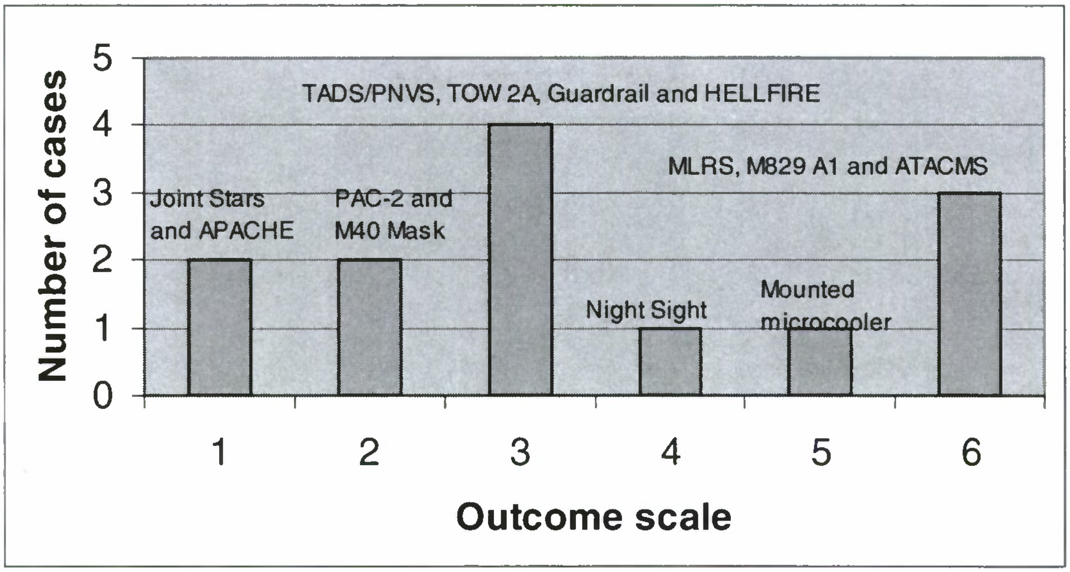

Figure 3.8 – System outcomes, scaled (O5-O10)

4. SPECIFIC FINDINGS

SCOPE

This chapter of Volume I reports the significant relationships that are found between the project outcome variables and other project factors, based on analysis of data gathered using the research questionnaires. As noted in Table 2.1 of the INTRODUCTION, the questionnaire was designed to acquire the views of those interviewed on a wide range of factors anticipated to have some influence on the system acquisition outcomes. For example, the maturity of (up to) three key technologies being integrated into each system was assessed with the intent to develop information that could be correlated with conclusions of a recent report from the General Accounting Office on technology readiness levels. With another GAO report in mind, systematic information on the approach to testing used in systems development was acquired. In addition, various questions were used to examine issues of the stability of project funding, user support, and system requirements

The study also considers a broad range of questions which the investigators believed important based both on prior experience and the results of related research. The issues addressed drew from the results of the LeanTEC project, a four year study of the development and transition of technology-dependent systems in the aerospace industry, supported by a cooperative research agreement between the U.S. Air Force Manufacturing Technology Office and The Boeing Company. The LeanTEC team guiding that research included representatives from six defense contractor firms, the Air Force, and university researchers. That research focused on projects applying new technology to production systems and the insertion of new technology in both civilian and military systems, involving technologies that ranged from better paints for titanium surfaces to combat systems electronics. It began with a year of unstructured interviews identifying key factors believed to impede or facilitate the development and transition of technology into production systems. Then a structured instrument was developed to ask veteran technical professionals consistent questions about projects they had worked on. Data were collected on just under 300 projects and the key results summarized here are interesting because of their similarity to the results from the present study.

Some 50 issues found to be important in the LeanTEC project were used in this present research, many using questions in exactly the same format. For example, a series of questions asked about the project team's composition during system development, the quality of its leadership, the continuity of its staffing, whether it was collocated, whether its members had worked together before, its cross-boundary work with production departments and suppliers including the joint use of prototyping, how early production personnel had been involved and the continuity of their involvement across the various stages of technology selection, development and transition to production. (See sections H, D and W in Appendix.) Several of these factors found important in the LeanTEC study emerged again as significant predictors of success for the Desert Storm programs included here, and are discussed in the following paragraphs.

TEAM CHARACTERISTICS AND PRACTICES

Team Leadership

In the prior LeanTEC field work, four characteristics of a team leader were mentioned as being more important than others. Some of those interviewed noted that it had been important that the team leader was good at resolving technical differences of the team when choices had to be made, and others made a related suggestion that the higher the technical competence of the leader, the better they were able to provide respected leadership. Another major leadership skill frequently mentioned was the ability to identify and deliver the resources the team needed in its work. A fourth leadership question asked whether team leaders had both engineering design and production experience, enabling them to work more effectively with both communities. In the subsequent quantitative phase of the LeanTEC work, project outcomes were found to be correlated with leadership.

Table 4.1
Leadership and Staffing Questions

Question*	Focus
D1: The team leader was good at resolving technical disagreements.	Technical leadership
D2: The team leader was good at getting necessary resources.	Resource leadership
D3: There was a lot of turn-over in team membership.	Staffing stability
D4: The team leader had both design & production experience.	Leader skill breadth
D5: The team leader had very high technical competence.	Leader technical skill
D6: Some key technical skills were not represented on the team itself.	Staffing practice
D8: Professionals were split across too many different tasks & teams.	Over-commitment
D10: Key members continued through pre-production planning and testing.	Staffing continuity

* Statements were posed for informants to provide answers varying from strongly agree to strongly disagree. "Strong" responses are treated as confident judgments.

These questions were repeated in the current investigation (Table 4.1), and similar results were found between project outcomes and leadership. Some of the strongest differences were linked to the leader's ability to deliver necessary resources (D2). Where the judgment had been made with confidence that the project leader had this skill, three of the six systems avoided late engineering changes, and the other three only required minor changes. When there was some reservation about the leader's ability to obtain resources for the program, it was found that none of the seven had avoided problems, and two had encountered significant late engineering changes (Table 4.2A.) A slightly stronger relationship was found between this leadership skill and staying within budget. Four of six projects with this type of leadership stayed within budget, and none badly exceeded budget. Where the leader could not be confidently judged to have skill getting resources, none met budget and four of seven significantly exceeded budget (Table 4.2B). Despite the small number of projects being studied, these differences are statistically significant and both could have only occurred by chance less than one time in a thousand ($p < .001$).

TABLE 4.2
Team Leadership and Program Outcomes

Late engineering changes after production had started? (O6)	A. Team leader good at getting necessary resources (D2) #	Other responses	#	Strongly agree
Significant	2	28.6%	0	0.0%
Minor changes	5	71.4%	3	50.0%
No, or very minor changes	0	0.0%	3	50.0%
Total	7	100.0%	6	100.0%

Kendall's Tau B = 0.614, significant at .001.

System met budget goals? (O8)	B. Team leader good at getting necessary resources (D2) #	Other responses	#	Strongly agree
Significantly exceeded budget.	4	57.1%	0	0.0%
Slightly exceeded budget.	3	42.9%	2	33.3%
Met budget.	0	0.0%	4	66.7%
Total	7	100.0%	6	100.0%

Kendall's Tau B = 0.742, Significant at .001.

A note on the statistical analysis. Analysis of variance and other statistical models typically used with this type of data must make strong assumptions about the level of measurement and the nature of the underlying distribution of the variables which do not seem warranted here. Further, the small number of cases makes impractical the statistical adjustments sometimes used to justify other approaches. Consequently the results are tested by the Tau B statistic designed for use when relating two ordinal variables, which is to say categorical variables that are not continuous but do show a consistent increase or decrease (of imprecise magnitude as from "minor changes" to "significant changes") from one category to the next.

To avoid inadvertently exaggerating the results, the actual number of cases is included in the tables to remind the reader of the small empirical base of the study. At the same time the Tau B statistic is used to show whether the results could have happened by chance. Thus in Table 4.1B above, the relationship between a leader being reported as good at getting resources and meeting budget goals is significant at the .001 level, meaning this distribution of cases could have only occurred by chance one time in a thousand, far less likely than the one chance in twenty commonly used in social science as a criterion for accepting that a relationship exists.

The Tau B statistic for two variables A and B is determined by first calculating how many cases would be properly assigned in a AxB table by chance given the separate distribution of their answers. That step is taken in turn for each cell of the table, such as the A_2B_1 cell (cases with the second value of A and first value of B) by using the number of cases with value A_2 divided by the total cases N, multiplied by the number of cases with a value of B_1. Summed across the cells this yields the number of cases which would be assigned correctly to cells in the AxB table by chance, leaving the remaining cases to be the errors that are expected if no relationship exists between A and B. (This method is similar to the Chi Square calculation of f_e.)

The Tau B value is the proportion of the total expected errors by chance that are reduced by the existence of an AB relationship. Given the assumption that a perfect AB ordinal relationship would allow accurate assignment of all cases (Tau B = 1.000), one asks what proportion of the errors which would have been expected by chance are reduced by using variable A to predict the distribution of B. Table 4.2B shows that .742 of errors expected by chance are explained by the presence of a relationship between resource leadership and meeting budgetary goals.

The Tau B for this and all following tables are computed for all values of the variables. However, the spread of a small number of cases in the full tables can make patterns more difficult to see. After the Tau B is calculated, the categories are collapsed for the convenience of the reader, so that for example one compares the strongly agree responses on D2 above with all other responses on D2 combined under "Other responses."

When other leadership skills are related to outcomes, there is further support for the importance of program leadership. The same leaders thought to be good at getting resources were generally those believed to have the ability to resolve technical differences effectively (D1), and not surprisingly some of the same relationships are found for the two types of leadership. In particular, the results are identical for the relationships between resource leadership and, separately, technical leadership with late engineering changes. Three of the six teams with leaders confidently judged able to sort out technical conflicts had no or only very minor late engineering changes, while none of the seven teams where the informants were less confidant that the team leader had this skill avoided late changes (Table 4.3). The relationship between resource leadership and late changes (not shown) is the same. The differences for both tables are highly unlikely (.001, or 1 in a thousand) to have happened by chance.

Table 4.3
Technical Leadership and Engineering Changes

	Leader good at resolving technical differences (D1)			
Late engineering changes after production had started? (O6)	#	Other responses	#	Strongly agree
Significant changes	2	28.6%	0	0.0%
Minor changes	5	71.4%	3	50.0%
No, or very minor changes	0	0.0%	3	50.0%
Total	7	100.0%	6	100.0%

Kendall's Tau B = 0.614, significant at .001

In previous research, some of those interviewed asserted that team leaders with both design and manufacturing experience were more effective because they could provide unique insights into problems and solutions. This survey consequently asked if project leaders had both kinds of experience (D4), and informants were able to make that judgment for 12 of the cases. The results are that this breadth of experience is not related to most of the outcomes, but it is found to be strongly related to meeting program budget. Three of the four cases with leaders with both design and production experience stayed within budget. Only one case with a leader not seen as having both types of experience is equally successful (Table 4.4).

Table 4.4
Team Leadership Experience and Meeting Budget

	Leader had both design and production experience (D4)			
System met budget goals? (O8)	#	Other responses	#	Strongly agree
Significantly exceeded budget	3	12.5%	1	0.0%
Slightly exceeded budget	4	50.0%	0	25.0%
Met budget	1	37.5%	3	75.0%
Total	8	100.0%	4	100.0%

Kendall's Tau B = 0.620, significant at .008 for N=12 cases (outcome data missing for one case).

A summary of the differences in the over-all effects of the four leadership characteristics can be provided by using an outcome metric which combines the six outcomes that were collected for each system. By asking for each project how many of most favorable levels are achieved across the six outcome questions being considered, a simple scale from 0 to 6 can be created. That is to say, a system that met its (1) technical, (2) program budget and (3) systems cost goals, was (4) completed on schedule, (5) had no late engineering changes, and (6) met expectations when deployed in the Desert Storm theater is scored with a 6. The actual 13 projects considered here vary from having achieved one to six of the desirable outcomes. (See GENERAL FINDINGS.)

When they were grouped by judgments made about the four leadership characteristics, the greatest difference is found between those programs where informants confidently reported that the leader was good at getting resources and at resolving technical differences (usually referring to the same leaders), and those that were not (Table 4.5). For example the seven projects with leaders less effective at getting resources are successful on just over an average of two (2.29) of six outcomes, while those programs with better leaders average close to five (4.83) successful outcomes, a difference that could have occurred by chance only 7 times in a thousand. Programs where leaders were capable of resolving technical differences were successful on 4.33 outcomes, and 2.71 outcomes when they were not, a difference that might be though meaningful but could have occurred roughly one time in 10 by chance. By contrast, technical skill alone is not related to a noticeable difference in the number of successful outcomes.

Table 4.5
Team Leader Capabilities and Team Performance
(Average number of successful outcomes and N)

Team leader capabilities	Other responses	Strongly Agree	Signif. at*
Good at getting resources (D2)	2.29 (7)	4.83 (6)	.007
Good at resolving technical differences (D1)	2.71 (7)	4.33 (6)	.105
Had both design and production experience (D4)	3.00 (8)	4.00 (4)	n.s.
Had very high technical competence (D5)	3.20 (5)	3.62 (8)	n.s.

*Significance of differences of means calculated here and in following tables using t test not assuming equal variances. Data not available on one case for D5.

Team Staffing Practices

In the earlier LeanTEC research, a central finding was that staffing as practiced in the aerospace industry was a critically important problem. The teams studied were generally established as cross-functional teams that were following the concept of concurrent engineering: The projects were staffed by integrated product development teams, they were composed of a mix of specialists believed necessary to applying a technology and transitioning it into production, and had authority/resources to call on departments and specialists for assistance. In the qualitative phase of the project experienced professionals suggested that while on paper their projects had met the formal definition of cross-functional

teams, in practice several problems were common. In particular, it was believed that some teams had encountered significant (sometimes fatal) problems because of inattention that resulted from assigning technical professionals to too many different teams and other responsibilities. Other complaints were that teams were expected to rely on key specialists not participating in the on-going discussions, leading to miscommunication and error, and that often teams were broken up and reassigned to new tasks before the transition into production was complete. Such practices were found to be among the strongest predictors of poor team performance.

In the present study, the same questions about staffing practices were repeated in the interviews on the Desert Storm cases used here (Items D3, D6, D8 and D10 in Table 4.1). All four staffing practices related to poor project performance to a limited degree, but the two items related to continuity were substantially more influential. The most important is concerned about continuity of individual engagement over time, or more simply, had there been substantial turn-over. When one compares the projects where the informants were confident that the projects did not experience substantial team turnover (Table 4.6A), five of the seven projects (71.4%) met the systems cost requirements set for the program, compared to only one in six of those where there was less confidence about the continuity of staffing. Turn-over also relates to lower performance against technical requirements, (Table 4.6B). For

Table 4.6
Continuity of Staffing Practices

A. Lot of turn-over in team membership (D3)

System met cost goals? (O7)	#	Other responses	#	Strongly disagree
Fell far short of cost goals.	1	16.7%	0	0.0%
Came close to cost goals.	4	66.6%	2	28.6%
Met or exceeded cost goals.	1	16.7%	5	71.4%
Total	6	100.0%	7	100.0%

Kendall's Tau B = 0.632, significant at .010.

B. Lot of turn-over in team membership (D3)

System met technical requirements? (O9)	#	Other responses	#	Strongly disagree
Fell short of meeting goals	4	66.7%	0	0.0%
Met or close to meeting goals.	2	33.3%	7	100.0%
Total	6	100.0%	7	100.0%

Kendall's Tau B = 0.729, significant at .001.

C. Key members continued through pre-production (D10)

System came in on budget? (O8)	#	Other responses	#	Strongly agree
Significantly exceeded budget.	4	66.7%	0	0.0%
Slightly exceeded budget.	2	33.3%	3	42.9%
Met budget.	0	0.0%	4	57.1%
Total	6	100.0%	7	100.0%

Kendall's Tau B = 0.742, Significant at .001.

seven projects where the respondents strongly disagreed that there had been turn-over, all seven had fully met their requirements, compared to two of six of those where the informants were less confident that staffing had been stable. (Note that the question was framed negatively, so strong disagreement is an assertion that there had been no consequential turn-over.) The statistical significance of these relationships, like many found in this study, is quite strong: the first relationships could have happened by chance only once in 100 times by chance, while the second at .001 could occur less than one time in a 1000 by chance.

Another strong difference is found when one looks to see how keeping the team together to facilitate transition to production relates to projects coming in on budget (Table 4.6C). Here the concern is whether there was continuing support from the team for the sometimes difficult problems that occur late in the transition to production and after production actually begins. Four of seven projects with key members continuing through pre-production met budget, compared to none of six where continuity did not seem to have been maintained. The only projects that were over a year late were those where informants had had some doubts about whether key members of the team stayed on into pre-production.

One can again look at the over-all effects of different staffing practices by looking at their relationships with the six point index of project success (Table 4.7). Turn-over and holding a core of the team together are related to a doubling of the success criteria that are met. Where the presence of turn-over is strongly denied, an average of 4.71 criteria were met compared to 2.00 for the other cases. Programs where key team members continued on through early production are seen to have met an average of 4.57 criteria, compared to 2.17. While the differences are not as great, having program teams which were over-committed is also related to some difference in outcomes. These results suggest that there were staffing practices in the development of Army systems in the period leading to Desert Storm that tolerated significant loss of continuity, and where there was a loss of continuity one finds much lower program performance.

Table 4.7
Team Staffing Practices and Project Performance
(Average number of successful outcomes and response N)

Team staffing practices	Other responses	Strongly disagree	Signif. at
A lot of turn-over among team members (D3)	2.00 (6)	4.71 (7)	.002
A lack of continuity into pre-production (D10)	2.17 (6)	4.57 (7)	.008
Team members assigned too many tasks (D8)	2.50 (6)	4.29 (7)	.067
Key specialties were not on the team (D6)	3.50 (8)	3.40 (5)	n.s.

Consistent with the LeanTEC research before the current study, staffing practices are a largely unrecognized source of substantial problems in development teams.

TESTING AND SIMULATION

Testing is a key process employed in weapons systems development to validate the progress being made in translating a concept into an actual product. Simulation is employed to both guide the choice of test conditions and to augment the testing process, since simulation allows for the estimation of component or system behavior over a much wider range of conditions than can be tested affordably. The results of tests are used to verify or "anchor" a simulation, so that it represents with adequate fidelity the behavior of a component or the system.

The General Accounting Office issued a recent report (GAO/NSIAD-00-199, July 2000, "A More Constructive Test Approach is Key to Better Weapon System Outcomes") in which the authors noted differences between what they observed in the testing approach employed by successful commercial firms and the approach employed in several major defense system development projects. The GAO report defined three levels of (integration) maturity that should be validated by testing during the development of a system. These are: 1. Components work individually, 2. Components work together as a system in a controlled setting, and 3. Components work together as a system in a realistic environment. The GAO authors argued that the more successful projects used an approach to testing which allowed reaching of the first two levels of integration maturity early in development. The GAO report noted that the earlier in development a (design) problem is discovered, the less expensive it is to fix. The report also described a dysfunctional phenomenon, "late cycle churn", wherein a significant problem is discovered late in development, presumably because of a faulty test process that defers key testing until very late in development. The GAO authors identified two principles as testing best practices: 1. Conduct the right validation events (tests) at the right time, and 2. Schedule challenging validation events early to expose weaknesses in the system design.

While based on a limited number of examples, this GAO report made a credible case that the testing approach used impacted project outcomes. Largely because of this, a series of questions (V1 to V15) were included in the research questionnaire in order to systematically acquire information on the testing and simulation processes employed in the cases studied in this research. It was expected that analysis of the answers to these questions might validate the arguments presented on the impact of the testing process employed on project outcomes.

It was anticipated that half or more of the fifteen testing and simulation questions might correlate with the various outcome measures. However, most of the projects are found to

Table 4.8
Test and Validation Questions

Question*	Focus
V9: Validation work used appropriately to improve system.	Testing utility
V11: Component & system maturity were validated at the right times.	Testing timing
V13: Validation events produced quality results.	Testing quality

* Statements were posed for informants to provide answers varying from strongly agree to strongly disagree. "Strong" responses are treated as confident judgments.

have uniformly conducted most of the test and validation activities being asked about. As a consequence there is little variation in the answers to allow them to be studied statistically, and one can only say no conclusion can be drawn through quantitative analysis. On testing issues not addressed in the discussion that follows here, the reader is referred to the attached case studies for such insights they might provide. (See, for example, the ATACMS case.) Where variation in the answers to the survey questions is found, two factors (V11 and V13) about the timing and quality of the testing are found to be significantly correlated with program outcomes.

Effective test and validation work requires that various events are timed to provide the best guidance to the system developers. As shown in Table 4.9, V11 ("Component and system maturity were validated at the right times in the program") correlated positively with the extent to which the system unit cost goals were achieved, with development budget performance, and with whether problems were encountered in the field during Desert Storm.

TABLE 4.9
Timing of Testing and Validation Events and Outcomes

	A. Component and system maturity were validated at the right times in the program (V11)			
System met cost goals? (O7)	#	Other responses	#	Strongly agree
Fell far short of cost goals.	1	16.7%	0	0.0%
Came close to cost goals.	4	66.6%	2	28.6%
Met or exceeded cost goals	1	16.7%	5	71.4%
Total	6	100.0%	7	100.0%

Kendall's Tau B = 0.606, significant at .002.

	B. Component and system maturity were validated at the right times in the program (V11)			
System met budget goals? (O8)	#	Other responses	#	Strongly agree
Significantly exceeded budget.	3	50.0%	1	14.2%
Slightly exceeded budget.	2	33.3%	3	42.9%
Met budget.	1	16.7%	3	42.9%
Total	6	100.0%	7	100.0%

Kendall's Tau B = 0.505, significant at .036.

	C. Component and system maturity were validated at the right times in the program (V11)			
Operational problems in the field? (O10)	#	Other responses	#	Strongly agree
Field problems limited effectiveness	4	66.7%	1	14.3%
Deployed at no loss of effectiveness	2	33.3%	4	85.7%
Exceeded expectations	0	0.0%	2	28.6%
Total	6	100.0%	7	100.0%

Kendall's Tau B = 0.505, significant at .006.

The strongest relationship is between the timing of the testing and meeting the system unit cost goals, with five of seven programs said to have correctly timed their testing meeting their goal. Only one of 6 that had not timed its testing well met its cost goals. (Table 4.9A.) A possible interpretation is that timely testing allows for equally timely tradeoffs to be made as design choices are made that influence production costs.

Timely testing also had a positive influence on the likelihood of good development budget performance and avoiding operational problems in the field. These findings seem intuitively correct. Late testing can provide critical information that forces corrective work, adding staff and other costs which might have been avoided had testing been performed earlier, a view consistent with the finding that three out of the four projects that met their budget had appropriately timed testing. Conversely, of the four projects that significantly exceeded budget, three of four were judged to have been less able to conduct testing at the right times (Table 4.9 B). Testing too early when changes are being made or too late for minor corrections would also be expected to run the risk of weaker field performance. Table 4.9C shows that there are five Desert Storm cases that encountered problems in the field that limited systems effectiveness, and four of these five are said to have not conducted testing at the best times.

The most significant correlations with outcomes for the other testing variable, V13 ("Most project validation events produced quality results"), to be discussed here are shown in Table 4.10. In Table 4.10A, two of the five cases in which the informants agreed strongly that the test (and simulation) program produced quality results encountered minimal changes, and the three that had changes had only minor ones. Where the quality of testing was more

TABLE 4.10
Quality of Testing and Validation and Engineering Changes

A. Most project validation events produced quality results (V13)

Late engineering changes after production had started? (O6)	#	Other responses	#	Strongly agree
Significant changes	2	25.0%	0	0.0%
Minor changes	5	62.5%	3	60.0%
None, almost none	1	12.5%	2	40.0%
Total	8	100.0%	5	100.0%

Kendall's Tau B = 0.532, significant at .009.

B. Most project validation events produced quality results (V13)

System met cost goals? (O7)	#	Other responses	#	Strongly agree
Fell far short of cost goals	1	12.5%	0	0.0%
Came close to cost goals	5	62.5%	1	20.0%
Met or exceeded cost goals	2	25.0%	4	80.0%
Total	8	100.0%	5	100.0%

Kendall's Tau B = 0.386, significant at .055.

doubtful, only one of eight avoided late changes during production and two of the seven that had this problem are the only cases to have significant late engineering changes.

Again looking at the programs where there are reservations about the quality of testing, one finds four out of five cases met their unit cost goals, and the fifth case came close. Only two of eight cases reporting less confidence in the testing quality met their cost goals, and one finds here the only case that fell far short of achieving its goal (Table 4.10B).

As was done in the preceding section on team characteristics and practices, and using the same method, Table 4.11 contains a summary outcome metric comparison for these testing and simulation variables. Here one sees how differences between strong agreement and other responses about appropriate timing (V11), quality of results (V13), and appropriate use was made of validation results (V9) relate to the average number of successful results on systems programs. Those systems which had timely testing and simulation events met on average 4.43 of the success criteria, compared to 2.33 for those which did not. The programs which had produced quality results were successful on 4.60 criteria, compared to 2.75 of those which did not. There is support here for the GAO conclusion that good testing programs are a key to project success, particularly as it regards conducting the right validation events at the right time.

Table 4.11
Testing and Validation Effectiveness and Project Performance
(Average number of successful outcomes and response N)

Types of testing and validation:	Other responses	Strongly Agree	Signif. at
Component and system maturity were validated at the right times in the program (V11)	2.33 (6)	4.43 (7)	.030
Most project validation events produced quality results (V13)	2.75 (8)	4.60 (5)	.069
Knowledge from validation work used consistently to improve components and system (V9)	3.60 (5)	3.38 (8)	n.s.

Significance of differences of means calculated using t test not assuming equal variances.

PROGRAM STABILITY

Previous reports on systems development issues have noted the importance of factors that influence the stability of system acquisition programs, with most attention having been paid to the impact of funding changes (typically, reductions). Accordingly, interview questions were included in this study to investigate several aspects of instability that might have impacted the systems being studied. Program funding uncertainty and changes, changes to the system requirements (e.g. changes to the threat the system was being designed to deal with) and changes in key TRADOC (or other user) personnel are all examined to see to what extent any or all of these "instabilities" impacted program outcomes. The table that follows contains the several questions which are used in the research categorized by type of stability or instability.

There was some a priori expectation that these various types of instability might not be independent. For example, funding instability might be linked to changes in requirements which occurred during the development or transition to production stages of the program. Or a relationship might be expected to exist between changes in key user personnel and changes in either funding or system requirements.

Table 4.12
Program Stability Questions

Question	Type of Instability
D11: There was often uncertainty about the future of project funding?*	Funding uncertainty
H1: At some point, was the project either slowed down or stopped and restarted? [No projects studied here had been stopped & started.]	Project slow-down
B2: Did cut-backs in project resources force changes/compromises?**	Funding cut-backs
F6: Were there changes in key TRADOC or other user personnel during development?***	Turn-over in user personnel
F7: How often were there changes in system requirements (e.g. threat) during development?	Change in requirements
W4: When (what stages) was there change in key TRADOC/user representatives?***	Change in user representation
W5: When did TRADOC/other users show strong support?***	Variation in user support
B13: Threat definition/other requirements changed during the project?**	Change in requirements
W6: When (during which stages) were there changes in the systems requirements?***	Change in system requirements
F5: How often did TRADOC/other user representatives show strong support during development?	Consistency of user support

* Statements were posed for informants to provide answers varying from strongly agree to strongly disagree. "Strong" responses are treated as confident judgments.
** Work effort from "None" to "Major" resulting from cut-backs.
*** Responses selected as many periods as applicable from the stages of planning; early, mid- and late development; and transition.

Funding

Financial uncertainty was common among these Desert Storm cases, with the informants of ten projects unwilling to disagree strongly on D11 that there had been no uncertainly over funding during the program. The presence of instability of support is also evident in H1, a question about the continuity of the project and whether projects had been slowed or stopped and started. While none of these projects were stopped and restarted, five of the 13 were reported to have been slowed down at some point. Interestingly there is some over-lap in projects that were slowed down and those suffering from financial uncertainty, but the size of that over-lap suggests that to a substantial degree they are separate factors.

When financial uncertainty was present, it appears to have had significant consequences for the Desert Storm development cases. All three of the projects which did not face financial uncertainly also avoided problems caused by cut-backs, while only two of 10 that faced some degree of financial uncertainty avoided those problems (Table 4.13A). When one looks at the projects that are reported to have been slowed, all five experienced problems due to financial cut-backs. By comparison, only three of eight that were not slowed also experienced problems due to cut-backs (Table 4.13B). While program slow-down may be caused by a variety of factors besides or in addition to budget cuts (see the following Requirements discussion), once slowed, programs seem to have financial problems. These results are of course not surprising. It is more that they provide reassurance that the informant judgments of the projects are consistent.

Table 4.13
Funding Instability and Forced Changes and Compromises

A. Often uncertainty about future of project funding? (D11)

Cut-backs forced changes (B2)	#	Other responses	#	Strongly disagree
Changes were forced by cut-backs	8	80.0%	0	0.0%
No changes forced by cut-backs	2	20.0%	3	100.0%
Total	10	100.0%	3	100.0%

Kendall's Tau B = 0.683, significant at .001

B. Was project slowed down? (H1)

Cut-backs forced changes (B2)	#	Slowed down	#	Kept on schedule
Changes were forced by cut-backs	5	100.0%	3	37.5%
No changes forced by cut-backs	0	0.0%	5	62.5%
Total	5	100.0%	8	100.0%

Kendall's Tau B = 0.688, significant at .001.

Staffing, Leadership and Funding Uncertainty. In the earlier LeanTEC study which has influenced this investigation, preliminary interviews suggested that there was some tie between funding stability and project performance. Veteran professionals in the aerospace industry had mentioned a number of projects that were weakened by perceptions that project

funds were limited or at risk. They suggested that when funding seemed threatened, development team engineers had a tendency to migrate to other, more stable projects causing turn-over. Being able to bill to multiple engineering charge numbers gives the individual substantial security and control over his work if the primary project encounters financial cut-backs or is cancelled. Another line of reasoning was that worry about continued funding led management and team leaders to cut back on staffing or otherwise reduce costs to stretch the project out. Whatever the reasons, informants were confident that they had seen a substantial number of projects where funding uncertainties had directly contributed to poor team performance.

Following those suggestions, the LeanTEC study included staffing questions in its collection of structured data, and the results confirmed that these views were correct. Informants were asked if they agreed or disagreed with the statement that there had often been uncertainty about the future of their projects funding, and their answers related to outcome questions similar to those used here. Staffing practices were found to be stronger for projects which informants disagreed with this statement indicating financial uncertainty had not been present (D11 in the current study). (Results from this earlier research are available by request.)

In the present study, the same pattern is found again. For the Desert Storm cases, when one looks in turn at how these questions about financial stability relate to other key factors, strong, negative relationships are found with staffing quality, effective testing, and to a lesser degree leadership. In particular, both financial uncertainty and cut-backs are found to relate strongly to turnover. When one examines the three projects where informants had strongly

Table 4.14
Funding Instability and Staff Turn-Over

		A. Uncertain Funding (D11)		
Member turned over? (D3)	#	Other responses	#	Strongly disagree
Other responses	6	60.0%	0	0.0%
Strongly disagree, no turnover	4	40.0 %	3	100.0%
Total	10	100.0%	3	100.0%

Kendall's Tau B = 0.667, significant at .001

		B. Cut-backs forced changes? (B2)		
Members turned over? (D3)	#	Forced changes	#	No
Other responses	6	75.0%	0	0.0%
Strongly disagree, no turnover	2	25.0%	5	100.0%
Total	8	100.0%	5	100.0%

Kendall's Tau B = 0.635, significant at .001

disagreed that funding was uncertain (D11), all three are reported to have had no turnover. For the remaining projects where funding was more uncertain, only four of ten avoided some suggestion of turnover (Table 4.14A). When one compares the projects which had not had

any compromises or changes forced by cut-backs with those that did, the results show that all five projects with no cut-backs also had no turnover. By contrast, only two of the eight projects with cut-backs on B2 avoided turnover (Table 4.14B).

Leadership. Views of resource leadership are linked to financial uncertainties. The three projects where the informants were confident that there was no doubt about funding are all found to be cases where they were sure that the leader was good at getting resources. Where funding uncertainties are judged to have been present, only three of ten confidently believed that the leader had been good with resources (Table 4.15A). Where the projects had been kept on schedule (H1), it would appear that leadership is held somewhat less responsible. Four of the five cases which were slowed down report that leadership was not good at getting resources. Where the projects moved ahead on schedule, informants in five out of eight cases credited the project leadership with skill for getting resources (Table 4.15B).

One should note that team leaders come in for some share of the blame for financial uncertainty whether or not they are responsible for the project's difficulties. The ambiguity of interpretation is that one cannot say whether the leaders are judged to be weaker at getting resources because of the cut-backs which may well have been driven by forces outside the leader's control or whether it is their lack of skill that led to the negative impact of the cut-backs, or a bit of both.

Table 4.15
Funding instability and Resource Leadership

	A. Often uncertainty about project funding? (D11)			
Leader good at resources? (D2)	#	Other responses	#	Strongly disagree
Other responses	7	70.0%	0	0.0%
Strongly agree	3	30.0%	3	100.0%
Total	10	100.0%	3	100.0%

Kendall's Tau B = 0.510, significant at .005

	B. Was project slowed down? (H1)			
Leader good at resources? (D2)	#	Slowed down	#	Kept on schedule
Other responses	4	80.0%	3	37.5%
Strongly agree	1	20.0%	5	62.5%
Total	5	100.0%	8	100.0%

Kendall's Tau B = 0.415, significant at .094.

Testing. Questions about project interruption, cut-backs and financial stability are also found to be related to testing as captured by V11, the appropriate timing of the testing used in the program. While the results suggest that uncertainty of funding is not strongly related to V11, the other factors considered here are. For the projects which were slowed during the life of the program (H1), four of five cases are also found not to have timed their testing activities well. For projects which were not stretched out, six of eight were reported as having conducted appropriate testing on V11 (Table 4.16A). Looking at B2, the degree to which cut-

backs caused changes and compromises, leads to a similar conclusion. Four out of five cases which did not have any changes forced by cut-backs also had appropriate testing. This compares with three of eight programs which were reported slowed (H1) which are reported confidently as having appropriately timed their testing.

Table 4.16
Funding Instability and Timing of Testing

A. Was project slowed down? (H1)

Appropriate timing of testing? (V11)	#	Slowed down	#	Kept on schedule
Other responses	4	80.0%	2	25.0%
Strongly agree	1	20.0%	6	75.0%
Total	5	100.0%	8	100.0%

Kendall's Tau B = 0.620, significant at .004.

B. Cut-backs forced changes (B2)

Appropriate timing of testing? (V11)	#	Occurred	#	None
Other responses	5	62.5%	1	20.0%
Strongly agree	3	37.5%	4	80.0%
Total	8	100.0%	5	100.0%

Kendall's Tau B = 0.610, significant at .002.

The over-all effect of financial-related problems on project performance is shown by again looking at the average number of successful outcomes which the cases reached (Table 4.17). The three cases which never had a problem with uncertain funding had an average of 5.67 successes out of a possible six; those that did face this uncertainly average 2.80. Cases which were never slowed (H1) averaged 4.50 success compared with 1.80 of those which were. Those cases which avoided changes caused by cut-backs average 4.13 compared to 2.40 for those which had that problem. It might be argued that financial uncertainty and cut-backs follow when projects encounter other difficulties, in which case these differences in averages

Table 4.17
Testing and Validation Effectiveness and Project Performance
(Average number of successful outcomes and response N)

	Other responses	Positive response*	Signif. at
Stability and funding:			
Uncertainty about project funding? (D11)	2.80 (10)	5.67 (3)	.001
Project ever slowed down? (H1)	1.80 (5)	4.50 (8)	.001
Cut-backs in resources forced changes (B2)	2.40 (5)	4.13 (8)	n.s.

*The positive responses are: D11, to disagree strong that funding was uncertain, H1, to say the project was never slowed or interrupted, and B2, cut-backs had no negative effects on the project.

would expect to be higher than those found for other factors, but there is little doubt that the presence of funding problems is strongly associated with poor development performance for these Desert Storm cases.

Given the small number of cases and the nature of the data, one must be cautious in asserting cause and effect relationships. On the other hand, experience suggests that stretching projects disrupts schedules, and that cut-backs and changes often lead to the need to repeat old test procedures or design new ones. That and the presence of turnover means that testing programs are sometimes being designed by different individuals from those that designed the system and supported its integration. Whatever the mechanisms, the general conclusion from these results is that instability and the loss of continuity seriously affect the quality of staffing and testing, which have been shown above to be in turn key predictors of weak program performance. Experience, the earlier LeanTEC research, and the pattern of these results are all consistent with an interpretation that uncertain funding and slowing and stretching projects impacts staffing, project leadership and testing, with a further impact through these factors to poor development performance.

Change in system requirements

As noted earlier, there is considerable anecdotal evidence suggesting that significant changes in systems requirements will adversely impact program outcomes, particularly schedule and/or cost. It is the existence of this evidence which makes experienced project managers extremely wary of permitting any changes in system requirements to occur. Sometimes, however, actions on the part of potential adversaries, referred to as "changes in the threat" can force the issue.

In the thirteen successful development cases studied, only three reported no change to system requirements once a system concept had evolved, and only four reported no change during the development phase of the project. A third of those cases which experienced change during development described that change as requiring "significant" or "major" effort, while the remaining two-thirds only required "minor" or "very minor" effort (see B13 in Appendix). Moreover, some cases experienced more than one instance of requirements change during development, with four of the nine describing encountering "several" or "many" changes (F7, Appendix). The remainder reported no change, or only one or two instances of change. The frequency of these changes seems to be at variance with the stability of perceptions of threats in the years prior to Desert Storm, a point to be revisited below.

When these factors are related to other variables, the results support the conventional wisdom that requirements changes are costly. Significant correlations were found between the three requirements change variables (B13, F7 and W6) and several of the outcome metrics, with F7 (which was designed to measure the frequency of change during the development phase) showing the greatest impact. None of the four projects which had several (3 cases) or many (1 case) requirements changes met their cost goals (Table 4.18A), and none of the four avoided late engineering changes (Table 4.18B). For those that had none or only one or two systems requirements changes, six of nine met their cost goals and three of

nine avoided late engineering changes. Weaker but similar differences are found for meeting budget goals (not shown).

Table 4.18
Changing Systems Requirements and Project Outcomes

	A. Frequency of changes in systems requirements (F7)			
System met cost goals? (O7)	#	Several, or many	#	None, or One or two
Fell far short of cost goals	1	25.0%	0	0.0%
Came close to cost goals	3	75.0%	3	33.3%
Met or exceeded cost goals	0	0.0%	6	66.7%
Total	4	100.0%	4	100.0%

Kendall's Tau B = 0.620, significant at .003.

	B. Frequency of changes in systems requirements (F7)			
Late engineering changes after production had started? (O6)	#	Several, or many times	#	None, or one or two
Significant changes	2	25.0%	0	0.0%
Minor changes	2	25.0%	6	66.7%
None, almost none	0	50.0%	3	33.3%
Total	4	100.0%	4	100.0%

Kendall's Tau B = 0.537, significant at .004.

The timing of systems requirement changes as well as their frequency seem to have an effect on project outcomes. Changes early in the development cycle are often easier to accommodate, but later changes may not be, particularly if they are have a major impact on the systems design. The interviews included questions which asked if systems requirements had changed at various stages of the project, including early, mid- and late in development. The strongest relationship found between these timing questions about systems requirements and outcomes is shown in Table 4.19. Four of the cases reported that systems requirements had

Table 4.19
Systems Requirements Changes in Mid-development and Field Performance

	Did systems requirements change mid-development? (W6d2)			
System performance in field? (O10)	#	Yes	#	No
Field problems limited effectiveness	3	75.0%	2	22.2%
Deployed at no loss of effectiveness	1	25.0%	5	55.6%
Exceeded expectations	0	0.0%	2	22.2%
Total	4	100.0%	9	100.0%

Kendall's Tau B = 0.485, significant at .027.

changed in mid-development, and of these in only one case does it appear that the system performed in the Desert Storm theater at a level which met expectations. By comparison, seven of nine cases which saw no changes in requirements in the middle of development met expectations.

One can see the over-all effects of the frequency and timing of systems requirements changes by looking at how many successful outcomes occurred on average for such cases (Table 4.20). A comparison of the average rates of success for changes in the three stage of development

Table 4.20
Stability of Systems Requirements and Project Performance
(Average number of successful outcomes and response N)

When systems requirements changed	Change occurred	No changes	Sig. at
Requirements changed in early development (W6d1)	3.75 (4)	3.33 (9)	n.s.
Requirements changed in mid-development (W6d2)	2.00 (4)	4.11 (9)	.013
Requirements changed in late development (W6d3)	3.33 (3)	3.50 (10)	n.s.
Frequency of requirements change	Several or many times	None, or or two times	Sig. at
Frequency of systems requirements change (F7)	1.50 (4)	4.33 (9)	.001

show that changes in mid-development are most strongly related to poor performance. Early changes do not seem to matter (3.75 and 3.33). The four projects that had systems requirements in mid-development only average 2.00 positive performance outcomes, compared to 4.11 average successes of those that did not. For these cases, changes in late development seem not to have an impact, and one can only speculate that perhaps changes this late are necessarily small. The four projects said to have seen systems requirements changes several or many times during development averaged only 1.50 successful outcomes, compared to 4.33 successes among those projects that had no, or only one or two, systems requirements changes. Both the frequency and timing of systems of changes in systems requirements are associated with poor systems development performance.

Change in key TRADOC (or other user) representative

The U.S. Army Training and Doctrine Command (TRADOC) is responsible for determining the requirements that Army materiel must meet in order to have utility on the battlefield. A senior TRADOC staff member (typically a colonel) is assigned to serve as the alter ego of the Project Manager insofar as interpreting these requirements as they are translated into system technical requirements during the acquisition process. This key individual may also play a critical role in preserving the planned funding for the system development by persuading more senior TRADOC leaders to strongly reaffirm the need for

the system when budget cuts are threatened or problems are encountered in the system development that increase cost or stretch schedule. As noted earlier, two variables (W4 and F4) were included in this study to attempt to assess the extent to which turnover in key TRADOC personnel might have influenced project outcomes. Similarly, two variables (W5 and F5) were included to examine the extent to which strong TRADOC support was important.

TRADOC Support. Strong user support is widely believed to be a critical factor for projects that wish to avoid the problems caused by the funding changes or related uncertainties discussed earlier in this section. Most of the projects enjoyed such strong support, however, that little can be said using the quantitative data. Eleven of the 13 success cases reported that TRADOC personnel showed strong support "many times" during the acquisition process, while the remaining two cases reported that TRADOC showed strong support "several times". Given this uniformity of response, the chances of obtaining significant correlation with outcomes is highly unlikely.

One can return to the qualitative responses for some insight. As mentioned above in the summary statements of critical problems, lack of user support was cited as a key reason for the two cases that did not make it successfully through development (See Table 3.3). The evidence available in this study supports the view that strong TRADOC support is important for systems to make it to the battlefield.

TRADOC Personnel Changes. Turnover in key TRADOC personnel was a common occurrence in the cases included in this study. Of the 12 cases where information on the timing of changes in TRADOC personnel was available, only two reported no change during the period of interest. The thirteenth successful case, which was not included in the statistics, reported that change in key TRADOC personnel occurred about every three years, but that the precise timing could not be recalled. This pattern of change is consistent with the military reassignment cycle.

As to whether TRADOC activity and change made a difference, significant relationships are found with several of the outcome metrics, notably O10, the extent to which the system met expectations when used on the battlefield during Desert Storm. Table 4.21 shows first the negative impact of change during early and late development on system performance on the battlefield (W4d1). There was change in key TRADOC personnel for five cases during early development, and four of those encountered operational field problems. Where there was no early TRADOC change, only one of seven projects was not as effective as expected. There is a suggestion in the data that TRADOC change in the later stage of development might also relate to operational problems in the field, but only change in this early stage of development is by itself statistically significant.

Another way to look at the possible effects of the turn-over of the Army's user representative is look at the consequences when there had been no reported changes in any stage of development (W4never). Only two cases, Night Sight and the M829A1 sabot, are found to have no TRADOC changes, and the relationship between no change at all and

operational performance is only marginally significant (although given the size of the Tau B statistic this is largely because only two cases are available in one of the comparison groups.)

Table 4.21
Continuity of Staffing Practices

A. Did TRADOC change during early development? (W4d1)

Operational problems in the field? (O10)	#	Yes	#	No
Field problems limited effectiveness	4	80.0%	1	14.3%
Deployed at no loss of effectiveness	1	20.0%	4	57.1%
Exceeded expectations	0	0.0%	2	28.6%
Total	5	100.0%	7	100.0%

Kendall's Tau B = 0.630, significant at .001. Data are available on 12 cases for W4d1.

B. Did TRADOC ever change after project start? (W4never)

Operational problems in the field? (O10)	#	Yes	#	Never
Field problems limited effectiveness	5	0.0%	0	0.0%
Deployed at no loss of effectiveness	6	0.0%	0	0.0%
Exceeded expectations	0	100.0%	2	100.0%
Total	11	100.0%	2	100.0%

Kendall's Tau B = 0.650, significant at .060.

It is striking that the two cases found with no TRADOC change are the same two cases that the informants felt exceeded operational expectations in the field (Table 4.21B).

Following the pattern of earlier sections and looking for a more general impact of TRADOC changes leads to the averages presented in Table 4.22. There is no consequential relationship of any kind between TRADOC change in mid- and late development and the scale of successful outcomes. Cases that experienced no TRADOC changes in early development are seen to be substantially more successful at an average of 4.29 successful outcomes, compared to an average of 2.40 for those that did have TRADOC changes at that time, although at .089 this could have happened by chance around one time in twelve. The conservative conclusion is that these results suggest that any direct, negative impacts of TRADOC changes are found largely in systems having less effectiveness in the field.

Table 4.22
TRADOC Changes and Project Performance
(Average number of successful outcomes and response N)

Timing of TRADOC changes	Changed	No change	Sig. at
TRADOC change during early development (W4d1)	2.40 (5)	4.29 (7)	.089
TRADOC change during middle development (W4d2)	3.86 (7)	3.00 (5)	n.s.
TRADOC change during late development (W4d3)	3.25 (4)	3.62 (8)	n.s.

It has been speculated that changes in key TRADOC personnel might be linked to a subsequent change in system requirements. This suggestion raises the possibility that TRADOC change could have adverse, indirect effects by somehow permitting changes in systems requirements which in their turn have a negative impact on project performance.

This study finds some support for that view when it examines the relationship between early to mid-development TRADOC changes and shifts in systems requirements. The results in Table 4.20 above suggest that the most damaging requirements changes are those that occur in mid-development. As a test of TRADOC change effects, one can aggregate TRADOC changes by asking how many of the earlier stages of development (W4s planning, W4d1 early and W4d2 mid-development) experienced turnover of key TRADOC personnel. In this way a number of 0 to 3 can be generated as a rough measure of continuing TRADOC turnover during the only stages when TRADOC change could logically have any influence on mid-development requirements changes. The results show that when there were no or only one stage that experienced TRADOC change, only one of seven projects experienced a mid-development change in systems requirements. When TRADOC changes occurred in two or all three of the early to middle project stages, three of five cases report that there were systems requirements changes during the middle of project development (Table 4.23). It would appear that TRADOC changes in the earlier and middle stages of projects are to some degree related to mid-development changes in systems requirements.

Table 4.23
Changing Systems Requirements and Key TRADOC Personnel

Did system requirements change during mid-development? (W6d2)	Number of early and mid stages TRADOC changes #	0 -1	#	2 -3
Yes	1	14.2%	3	60.0%
No	6	85.8%	2	40.0%
Total	7	100.0%	5	100.0%

Kendall's Tau B = -0.505, significant at 0.017.

DEALING WITH PROBLEMS

It is an often quoted truism that "bad news does not improve with age"; this seems to be particularly so when dealing with problems encountered in complex defense acquisition programs, since the flexibility to deal with problems by adjusting the design diminishes rapidly as time passes. Two questions were asked of those interviewed in this study in an attempt to determine whether problem communication delays had influenced program outcomes. Question D12 asked if the project team was reluctant to share concerns with the government Project Manager, while question D19 asked if the government Program Manager

was reluctant to share problems with Army leaders. The histogram in Figure 4.1 depicts the distribution of answers obtained for the 13 successful cases.

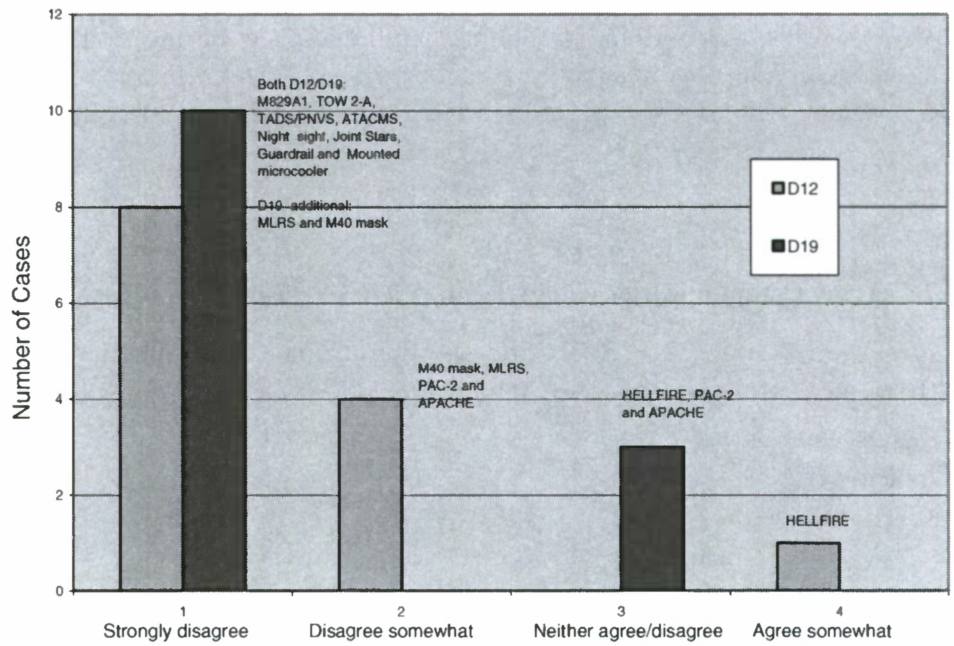

Figure 4.1 - Responses to D12 and D19

Note that in many of the cases studied, respondents strongly disagreed with the premise that there was any reluctance to disclose problems to either the government project manager

TABLE 4.24

Communications and Systems Operations Problems

A. Contractor-Army communication difficulties (D12)

Operational problems in the field?	#	Other responses	#	Strongly disagree
Field problems limited effectiveness	3	60.0%	2	25.0%
Deployed at no loss of effectiveness	2	40.0%	4	50.0%
Exceeded expectations	0	0.0%	2	25.0%
Total	5	100.0%	8	100.0%

Kendall's Tau B = 0.418, significant at .055.

B. Intra-Army communication difficulties (D19)

Operational problems in the field?	#	Other responses	#	Strongly disagree
Field problems limited effectiveness	3	100.0%	2	20.0%
Deployed at no loss of effectiveness	0	0.0%	6	60.0%
Exceeded expectations	0	0.0%	2	20.0%
Total	3	100.0%	10	100.0%

Kendall's Tau B = 0.608, significant at .010.

or to the Army senior leadership. Perhaps because of the relatively few cases where any reluctance to disclose problems in a timely way is found, few significant relationships between D12 and D19 and the outcome variables were found. The correlation of these two variables with O10, the extent to which the system met expectations when used on the battlefield, is shown in Table 4.24. Here, the impact of reluctance of the contractor to communicate with the Army leadership shows a somewhat weaker correlation with the occurrence of field problems than does reluctance of the Program Manager to communicate with Army leaders .

TABLE 4.25
Contractor-Army Communications and Organizational Resistance to Ideas

A. Contractor-Army communication difficulties (D12)

Impact of Army lab/center resistance (B11)	#	Other responses	#	Strongly disagree
Project spent effort on problem	4	80.0%	1	12.5%
Was not a problem	1	20.0%	7	87.5%
Total	5	100.0%	8	100.0%

Kendall's Tau B = 0.622, significant at .001.

B. Contractor-Army communication difficulties (D12)

Impact of Project Office resistance (B12)	#	Other responses	#	Strongly disagree
Project spent effort on problem	4	80.0%	3	37.5%
Was not a problem	1	20.0%	5	62.5%
Total	5	100.0%	8	100.0%

Kendall's Tau B = 0.574, significant at .010.

TABLE 4.26
Intra-Army Communications and Organizational Resistance to Ideas

A. Intra-Army communication difficulties (D19)

Impact of Army labs/centers resistance (B11)	#	Other responses	#	Strongly disagree
Project spent effort on problem	3	100.0%	2	20.0%
Was not a problem	0	0.0%	8	80.0%
Total	3	100.0%	10	100.0%

Kendall's Tau B = 0.619, significant at .017.

B. Intra-Army communication difficulties (D19)

Impact of Project Office resistance (B12)	#	Other responses	#	Strongly disagree
Project spent effort on problem	3	100.0%	4	40.0%
Was not a problem	0	0.0%	6	60.0%
Total	3	100.0%	10	100.0%

Kendall's Tau B = 0.571, significant at .016.

Relationships are also found between these reluctance communication variables and variables B11 and B12 which were designed to assess any costs associated with resistance to project team ideas and approaches. B11 queried about resistance on the part of Army science and technology organizations, while the similarly worded B12 asked about resistance on the part of Army Project Offices. Tables 4.25 and 4.26 show these relationships; the statistical significance is sufficiently strong that it is unlikely that they occurred by chance.

TECHNOLOGY READINESS

As noted earlier, this study was influenced by a 1999 U.S. General Accounting Office report, "Better Management of Technology Development Can Improve Weapon System Outcomes," and the recognition that Technology Readiness Levels are being widely used in management of systems development for NASA and the Department of Defense (DoD). The GAO investigators used a similar approach to the current study of asking comparative questions of a small set of projects, and drew interesting conclusions about the importance of technology maturity. Using the concept of Technology Readiness Levels (TRL) measured by a scale developed by NASA, the study assessed the impact of technology maturity on the performance of DoD systems development projects. Figure 4.2 below describes three of the nine TRLs used by DoD.

Referring to the use of TRLs in private industry, a principal finding of the GAO report was that DoD is inclined to start development with technologies at a lower readiness levels than general industry practice, and that the acceptance by DoD of technologies at TRLs lower than 5 contribute significantly to cost growth and schedule slippage. (See Figure 4.2.)

Figure 4.2

Technology Readiness Levels 4, 5 and 6*

4. **Component and/or bread board validation in lab environment.** Basic technological components are integrated to establish that pieces will work together, e.g., integration of ad hoc parts in lab. This is relatively "low fidelity" compared to the eventual system.

5. **Components and/or bread board validation in relevant environment**. Fidelity of breadboard technology is significantly increased. Basic components integrated with reasonably realistic supporting elements so the technology can be tested in a simulated environment. Examples include "high fidelity" laboratory integration of components.

6. **System/subsystem model or prototype demonstrated in a relevant environment.** Representative model or prototype system, which is well beyond the breadboard tested for TRL 5, tested in a relevant environment. Represents a major step up in a technology's demonstrated readiness. Examples include testing a prototype in a high fidelity laboratory environment or in a simulated operational environment.

*The full scale is found at the end of the APPENDIX.

When this GAO report was brought to the attention of the earlier LeanTEC project members, that research team was mounting a second, smaller wave of data collection to include more projects to test in detail some of the initial results. This set of projects consisted almost entirely of electronics projects introducing technology into components for larger systems. A decision was taken to add questions about technology readiness for each project being studied, asking informants to estimate the TRL levels when project planning began, when development began, when the transitional stage began to move the technology into a production system, and at the point when production began. The average TRLs found for these projects were 4.60 at the start of development, and still only at an average of 6.19 at the conclusion of development when they were accepted for transition into production. These results lead one to wonder if the GAO was accurate in suggesting that industry is generally more conservative and does not start projects with TRLs less than five.

There was some evidence that the LeanTEC TRL levels had an effect on project outcomes. The strongest relationship is found between the technology at its first consideration when planning had started and whether the project was successful in getting the system under development into production. Since the stages of systems production for these electronics and software projects all included a step of Low Rate Initial Production (LRIP), the outcomes of interest here were whether the system went into LRIP, and then whether it moved further by being transitioned as intended into full production or only a partial or modified use of the technology. As shown in Table 4.27, for the 39 LeanTEC cases, the five projects with TRLs at 5 or 6 were all brought to full production. There is a roughly linear increase from a single project that started with a TRL of one which failed, to a success rate of 53.3% at TRL 2, 62.5% at TRL 3, 80.0% at TRL 4, and finally a success rate of 100.0% for projects which were at a TRL of 5 or 6 at the time the project was planned.

Table 4.27
Technology Readiness Levels and Production Success
LeanTEC Aerospace Project

Did system go into production?	Technology Readiness Levels at Start of Planning									
	#	1	#	2	#	3	#	4	#	5 or 6
Abandoned, or shelved	1	100%	3	20.0%	0	0.0%	0	0.0%	0	0.0%
Stopped after LRIP	0	0.0%	3	20.0%	2	25.0%	2	20.0%	0	0.0%
Production of parts/ideas	0	0.0%	1	6.7%	1	12.5%	0	0/0%	0	0.0%
Reached full production	0	0.0%	8	53.3%	5	62.5%	8	80.0%	5	100.0%
Total	1	100.0%	15	100.0%	8	100.0%	10	100.0%	5	100.0%

Kendall's Tau B = .366, significant at .004 level (N=39)

To study TRL effects on other project outcomes, one must set aside the projects that did not move beyond LRIP, and focus on the quality of the project performance for teams that achieved some degree or full transition to production. Or said another way, one can only

study the effects of TRLs on project delay, late engineering changes or other outcomes for projects which were sufficiently successful that they reached a stage where these outcomes could be judged.

In this subset of cases from the LeanTEC project, TRL levels at the time planning started, at the time development started and at the time that a transition began to production generally do not relate to these project outcomes, contrary to the expectations that follow from the GAO report emphasizing TRL at development. The exceptions are that project TRLs at the time the projects actually entered production related to two outcomes (Table 4.28A). Not surprisingly all ten projects with TRLs still at six or seven when production actually started had late engineering changes, while roughly half of the projects with TRLs of 8 (45.5%) and 9

Table 4.28
TRL at Start of Production and Quality of Project Performance
LeanTEC Aerospace Project

A. Technology Readiness Level at Production

Late engineering changes	#	7	#	8	#	9
Significant late changes	1	12.5%	0	0.0%	1	11.1%
Minor late engineering changes	7	87.5%	4	44.4%	3	33.3%
None or almost no late changes	0	0.0%	5	45.5%	5	55.6%
Total	8	100.0%	9	100.0%	9	100.0%

Kendall's Tau B = .370, significant at .026.

B. Technology Readiness Level at Production

Met technical requirements	#	7	#	8	#	9
Close to meeting requirements	4	44.4%	2	16.7%	1	11.1%
Met or exceeded requirements.	5	55.6%	10	83.3%	8	88.9%
Total	7	100.0%	10	100.0%	9	100.0%

Kendall's Tau B = .366, significant at .043.

(55.6%) had no consequential late changes after that point. Higher TRLs also relate to projects meeting their technical requirements (Table 4.28B) with over 80% of the TRL 8 and TRL 9 projects meeting or exceeding their technical requirements; 55.6% of the TRL 7 projects meeting requirements, and the one TRL 6 project failing to meet requirements. The lack of late changes may mean that there is no need to compromise on requirements in order to get a system into full production. Also, no relationship was found between the TRLs at development start and at transition to production and the summed number of successful outcomes achieved by these LeanTec projects.

The LeanTEC results support the GAO finding that technology readiness matters, and that the NASA TRL scale can be used by knowledgeable professionals to assess the readiness of technical programs. The clear conclusion is that planning stage TRLs helps predict whether a

project will reach production. There was no support, however, for the view that TRL maturity specifically at the start of development predicts project performance for either the separate or the over-all number of outcomes. The only other relationships found with technology readiness are relationships between TRLs at the beginning of production and late engineering changes and meeting technical requirements, findings that would be surprising if they were not present.

Desert Storm TRLs. These results seem to contradict conventional wisdom about the use of TRLs, and led the present Desert Storm study to include questions about technology readiness. The Desert Storm informants on each case were asked to classify their systems on both the TRL for the systems as a whole, and for up to three key technologies involved in component subsystems. An immediate result is that it is evident that the systems TRLs of the 13 Desert Storm programs that reached production and eventually field deployment are remarkably similar in readiness to the LeanTEC mix of military and civilian systems (Table 4.29). Looking at the TRL scores in general, one can see the average TRLs of the Desert Storm and LeanTEC projects are similar. At the beginning of development the LeanTEC projects that had reached full production (for better comparability to the 13 Desert Storm projects) had an average TRL of 4.77. This average matches exactly the average TRL of 4.77 for the 13 Desert Storm systems included in our study that were produced and deployed for theater operations.*

Table 4.29
A Comparison of (Mean)Technology Readiness Levels in the LeanTEC and Desert Storm Projects

	Suggested GAO TRL benchmark	LeanTEC[a] project TRL	Desert Storm TRLs Integrated systems	Desert Storm TRLs Average technology	Desert Storm TRLs Lowest technology
At start of development	5.00	4.77	4.77	4.88	3.74
At start of transition to production		6.33	---[b]	7.87	7.67

[a] For comparability, these projects only include those that went beyond LRIP.

[b] Not ascertained in the study.

Given that the LeanTEC research focused on projects that were smaller, and often only involved a single new technology that was put into aerospace systems, a better comparison might be between the LeanTEC results and the TRLs of the component technologies that went

* Because of the importance of production problems in prior research, the investigators developed a Production Readiness Level (PRL) scale to see if that rather than general technology readiness is a key explanatory factor. The scale used is provided in the APPENDIX. The results are inconclusive, but do not contradict any of the findings here.

into the Desert Storm systems rather than the systems as a whole. The TRL ratings of the 36 component technologies at the start of development and the beginning of transition into production are shown in Table 4.29. The results suggest that the technologies being applied in the LeanTEC projects and Desert Storm programs were in a quite similar state of readiness whether one considers the TRL of the system as a whole, or the technology components.

A sceptic at this point might note that it is not the average but the least mature of the technologies which is most important in anticipating problems in development. For that reason the study also identified for each case the lowest technology level among the component technologies rated for that system, finding to an average TRL of 3.74 for the lowest technology TRL for each system.

Based on these largely 1990s electronics projects for aerospace systems from the LeanTEC study and the more complex systems developed for deployment in Desert Storm, the complex military systems developed for Desert Storm were not that different on the maturity of the technologies being adapted by industry, and roughly half of both sets fall below the GAO report suggested TRL benchmark of five, defined as, "Components and/or breadboard validation in relevant environment". If there are any differences, it is that the least mature component technologies in the Desert Storm systems have a rating well below the TRLs of both the GAO standard and the LeanTEC projects. One should note, however, that the gap was closed by the end of the development phase, and both the average and the least mature Desert Storm component technologies are seen to be higher than those of the LeanTEC projects (Table 4.29, row 2).

When the relationship between TRL levels and project outcomes are analyzed, the first finding is that -- like the LeanTEC result -- TRLs do predict differences in the basic result of achieving production status (Table 4.30). Looking at the Desert Storm cases in this study that completed development, all seven systems which had a TRL level of 5 or higher avoided all but very minor engineering changes in the transition to production, compared to four of six of those with lower TRLs, providing weak (very marginally statistically significant) support for the general view of the 1999 GAO report that TRLs are important. Notably this is the same relationship found in the LeanTEC projects.

Table 4.30
Technology Readiness Levels and Early Program Outcomes

	Technology readiness at start of development (TRL)			
Additional changes during transition to production started? (O2)	#	2-4	#	5-7
Significant changes	2	33.3%	0	0.0%
None or minor changes	4	66.7%	7	100.0%
Total	6	100.0%	7	100.0%

Kendall's Tau B = 0.447, not significant (.062).

When the TRL ratings are related to project outcomes for the 13 successful systems that moved through production and into the field, the projects with higher TRLs at the beginning of development were not that different than those with TRLs less than 5 on several outcomes. They are roughly similar in whether they were late, on budget, met their technical requirements and cost goals, and avoided late engineering changes. There is, however, a strong relationship between TRLs at the start of development and field performance (O10), but surprisingly they are in the opposite direction from that which is reasonably expected. All six of the programs that had TRLs of less than 5 either met or exceeded operational expectations when deployed in the field while only two of seven programs with a TRL of 5 or higher met that standard. (See Table 4.31A)

Table 4.31
Technology Readiness Levels and Program Outcomes

A. Technology readiness at start of transition stage
N=13

System performance in field? (O10)	#	2-4	#	5-7
Field problems limited effectiveness	0	0.0%	5	71.4%
Deployed at no loss of effectiveness	5	83.3%	1	14.3%
Exceeded expectations	1	16.7%	1	14.3%
Total	6	100.0%	7	100.0%

Kendall's Tau B = -0.551, significant at .001

B. Technology readiness at start of development
N=9 after four cases with mid-development requirements changes are removed.

System performance in field? (O10)	#	2-4	#	5-7
Field problems limited effectiveness	0	0.0%	2	50.0%
Deployed at no loss of effectiveness	4	80.0%	1	25.0%
Exceeded expectations	1	20.0%	1	25.0%
Total	5	100.0%	4	100.0%

Kendall's Tau B = -0.403, not significant (.135).

Analysis was conducted to see if (or which) other predictors of poor program performance related to TRL ratings that might account for this result. One observation is that by chance, the four projects that were most disrupted by systems requirements changes had high TRLs. When one looks at the timing of those changes, four projects experienced requirements changes in mid-development.

Given the importance of requirements stability reported above, an effort is then made to separate the effect of high TRLs from changing technical requirements. One can drop the four cases where changing technical requirements occurred in mid- development and look only at the remaining nine cases, setting aside some of the effects of changing requirements. The result is that the negative relationship between technology maturity and meeting

expectations in the field substantially drops. (See Table 4.31B.) Note that statistical significance will usually decrease by definition when the number of cases is reduced; the comparison is between the size of the two Tau B relationships.) In this case, the chance distribution of projects with unstable requirements seems to explain some of the negative relationship, although the negative relationship remains an anomaly. For our purposes here, however, the results are contrary to those predicted by the GAO study, and it seems safe to conclude that there is no evidence here that technology readiness at the start of development relates to better systems performance in the field.

Both the earlier LeanTEC research on smaller electronics projects and the Desert Storm cases on small and quite large systems programs lead to the conclusion that technology maturity is not a general predictor of program success. The evidence in both studies supports the belief that technology maturity predicts whether or not a project is successful to the point of being ready to transition to production. It might be assumed that those that cannot be ready for acceptance into production transition are terminated and in that sense TRL play a very important role. However, beyond that point the systems development processes appear to be able to compensate for residual technology weaknesses. For these Desert Storm systems that reached production and were deployed in Desert Storm operations, higher technology readiness does not predict delay, or failure to meet budget, technical requirements, cost goals or expectations for performance in the field.

5. CONCLUSIONS /RECOMMENDATIONS

After noting some important limitations of this study, the focus in this concluding chapter is on the single, encompassing theme which integrates the results of this research. The most important limitation is that while this study has proven to be remarkably robust in determining important predictive factors, it is inherently limited in that it cannot be said to have any implications for factors which may be important, but which are not able to be discussed for lack of numbers of cases and variation in key variables. What can be stated with confidence as a central conclusion is that a lack of program stability, whether manifested in changes in requirements, early turnover in key project staff members, or changes in program funding, is a key predictor of poorer outcomes in systems development. Implications of this conclusion for the management of military system development programs are interesting, and are briefly discussed.

Limitations

A study such as this one relies on the presence of variation in the factors that are studied. In instances where all, or alternatively, none, of the system development programs engaged in an organizational practice or experienced an external influence, this factor can only be inferred by reading the case studies in depth to find qualitative judgments made about its importance. Given the sheer diversity of most organizational practice over time, this problem is often a minor concern because so much variation is typically present. When one is limited to the data from 13 cases, as was the case for most of the analyses, this limitation can be more severe.

As an illustration of this problem, all 13 cases reported that testing had been done of the components working together in a controlled setting (V4). Going further, such integrated component testing was conducted in these cases by several organizations: in twelve cases by the prime contractor, in eight cases by suppliers, in nine cases by Army laboratories and centers, and in five other organizations. One cannot isolate the effects of integrated testing in general without variation, and separating the effects of which organizations did the testing with a small number of cases is problematical. In contrast, where there was variation on other questions on the quality and timing of testing, the results lent strong support to the importance of testing.

When a similar problem appears without an alternative question, there is little one can say. While it is widely accepted (and certainly intuitive) that lack of management support could adversely affect the outcome of a development project, since twelve of the 13 cases report confidently they were a management priority (D15) at the prime contractor (and the thirteenth agreed somewhat), there is insufficient variation here to contrast cases that did and did not have management support to do quantitative analysis. Without any alternative questions getting at the same issue, the research lacks evidence for or against the importance of those factors which could be significant determinants of relative project performance.

Another limitation resulting from the small number of cases is that it takes a very large effect to persist through the random measurement error which is always important in studies such as these. As noted previously, every variable used in this study has been identified as important (at least anecdotally) in one or more projects in the past, and nothing here implies that they are not influential in any of the projects studied here. The best way of stating the utility of the results which are reported in the preceding chapters is that they identify elements of program characteristics and activities that were generally important in explaining the effective development of the systems studied. Other factors may have been critically important in one or two cases, or somewhat important in many of them, and those effects would not have been captured in the statistical analysis. The reader is asked to turn to the qualitative case studies in Volume II of this report for analysis which may provide insight on these other determinants of success.

What can be said is that the quantitative analysis identifies the driving factors that were important in most of the cases, warranting attention to see if the problems continue today.

Central conclusion

Several of the statistically significant relationships discussed in Chapter 4 involve factors that are related to the stability of the program. When key members of the project team left the program too early, project outcome suffered. This turn-over may have been motivated by uncertainty or cut-backs in project funding; these funding instabilities themselves also correlated negatively with project outcomes. Further, both project funding cutbacks and project team turn-over negatively correlated with the quality of the testing program and the timeliness of key test events. These two attributes of the testing program also had the strongest correlation with project outcomes. In addition, changes in systems requirements during development correlated with poor project cost performance, and, finally, turn-over in key user representative personnel correlated negatively with system performance in the field.

Taken together, these several relationships strongly suggest that stability of program resources and objectives is a very powerful influence on the relative success of the project. Certainly, as has been noted, there is a wealth of anecdotal evidence that suggests that this should be the case. In reflecting on this array of instabilities that could impact a system development, it became clear that they had at least one thing in common. That is, the longer a system stayed in development, the greater chance it had to experience one or more of these program destabilizing events. Or, stated another way, shorter system development cycles should result in better project outcomes.

This hypothesis was tested by examining the correlation between the system development durations tabulated in Table 3.1 and the aggregate outcome scale described in Chapter 3 (and shown in Figure 3.8). This resulted in a strong correlation, with a Pearson's r statistic of 0.688, significant at 0.009. This can be interpreted as indicating that almost half (r squared = 0.473) of the outcome scale variation can be explained by development duration, with only nine chances in a thousand that the correlation is a random occurrence. A sensitivity analysis was performed to examine whether the data uncertainty noted in Table 3.1, which for several of the systems might have resulted in as much as six months more development duration,

impacted the strength of the correlation. Increasing the duration of these uncertain cases resulted in an almost identical value for the r statistic as calculated for the nominal durations and did not change the significance value adversely.

Consistent with this discussion, Table 5.1 also shows that a strong relationship exists between the previously used aggregate outcome scales and development duration.

Table 5.1
Length of Project Development and Project Performance
(Average number of successful outcomes)

	Over 3 years	Three years or less	Sig. at
Length of development	2.00	4.71	.002

Therefore, a central conclusion from this study is that shorter development cycle times favorably correlate with key project outcome variables, largely by minimizing the exposure of the project to destabilizing influences which have also been shown to correlate negatively with these same outcome variables.

The defense acquisition community has long recognized that lengthy systems development times are disadvantageous. Sometimes the associated negatives have been phrased in program instability terms and this study certainly provides a strong empirical support for those who hold these beliefs. Over the years a number of initiatives have been attempted to shorten development cycles, with limited success where complex systems were involved. The current approach is referred to as "spiral development"; its basic concept is to get a useful, if limited, capability in the field quickly and then introduce additional technology-based capabilities through further "spirals" of development. This approach appears to be in keeping with the implications of this study's central conclusion.

Consider the variety of advantages that occur with a shorter development cycle, in terms of the significant correlations identified in this research. Table 5.2 was developed to provide some insight on this issue; it displays the destabilizing influence, along with the implications associated with the length of the development cycle. Note that, in terms of minimizing the likelihood of destabilizing influences, shorter is clearly better. One could argue that planning for a development duration of three years or less would be prudent. Indeed, when one compares the development duration data in Table 3.3 with the outcome scale data in Figure 3.8, all the cases with development durations not exceeding 37 months attained ratings of three or higher on the outcome scale, and perhaps more striking, only cases with these shorter durations achieved an outcome scale value of four or higher.

Variable	Timing Implications
1. Reductions in project funding	Potential for change in administration every 48 months; typical turn-over in key military leaders occurs every 24-36 months. Potential change in key Congress positions every 24 months; likelihood increases with development duration
2. Uncertainty in project funding	Potential for change in administration every 48 months; typical turn-over in key military leaders occurs every 24-36 months. . Potential change in key Congress positions every 24 months; likelihood increases with development duration.
3. Change in system requirements	Changes in the threat environment occur unpredictably, but become more likely with longer development durations. Changes in doctrine and system requirements follow a similar pattern.
4. Change in key user representatives	Typical turn-over in such key military positions occurs every ~36 months
5. Change in key project team members	Typical turn-over in military acquisition positions occurs every ~36 months. Longer development durations present more opportunities for career moves on the part of key civilian team members

Table 5.2 – Destabilizing influences

Whether or not a change to selecting projects with shorter development times is made, the Army could do more to stabilize the guidance and resources given to both shorter and longer development projects. Acting alone, the Army could do more to map rotating personnel assignments and other sources of TRADOC change to project development cycles. It could eliminate all but the most critically important changes in systems requirements once projects move beyond early development since it appears that, as widely believed, such changes will almost certainly hurt project performance. Through contracts and informal management practices, the Army could work with its contractors to provide better continuity of development project staffing.

APPENDIX
Questionnaire and Frequency of Responses

This section includes the instructions and questions given to the informants on 15 selected cases, typically the project leader at the prime contractor for the Army system and the Army Project Manager. As described in the INTRODUCTION these were then used by students to guide their own study of a particular case, including subsequent interviews of cooperating prime managers and Army project officers. As a final step, the students wrote a detailed narrative case study and completed a master questionnaire, reconciling any differences between the answers of the informants. The master questionnaires were then used in the quantitative analysis.

Instead of spaces and blanks to record answers, the questionnaire here offers the results on the 13 cases that led to deployed systems. While in total 15 case studies were conducted, those about systems deployed in the field for Desert Storm operations are at the center of the analysis. The narrative cases on the two failures attached to this report contain useful lessons that should not be ignored, but because they did not reach field deployment, many of the outcomes asked for like meeting costs goals or field performance are not applicable. It was thought that results for the more successful 13 would be more useful for the reader.

In one section of the survey, questions were asked about the technologies that the new systems relied upon, with a request for up to three responses. A table of the responses is provided, and the reader should note that the informants listed three constituent technologies for ten systems, and two for the other three successful systems, for a total of 36 technologies. Then questions were asked about these separate technologies, and the results are reported for a base of 36 technologies, not 13 systems.

Last, there are some questions that reached a level of detail that neither the informants nor the researcher writing the case could provide an answer. The authors have gone back to the writers to check on the missing answers and in a few cases could elicit what were believed to be reliable answers for the survey instrument. When this checking was unable to arrive at a confident answer, the data was left as missing. Consequently the reader will find some reported results summing to data on only 12 or rarely 11 cases. Where it was believed to be useful, both the raw frequencies and the percentages are reported, and in some cases those percentages are calculated on the corresponding base of 13 or the lesser number where there were missing data.

Desert Storm Case Study Checklist: Lessons for Technology Management

The U.S. Army Materiel Command is supporting a hindsight study of how technologies were developed, integrated into systems, and produced in the years leading up to Desert Storm, the last large-scale deployment of U.S. military force. It is believed that in the years leading up to that conflict, there were both successful and unsuccessful applications of technology to military systems that contain lessons for future defense technology development. The study can be done now because the intervening years allow more objectivity, and allow open examination of what were once classified projects. The study must be done now because many of the men and women responsible for the development and eventual fielding of those systems in Gulf region are retiring, taking with them important knowledge that we believe should be captured and codified into practical lessons for the future.

Our method began with a list of military systems including both successes and failures judged to be broadly representative of the systems that were under development in the years prior to Desert Storm. Then experienced students (such as those found at senior military schools and mid-career management programs) are being asked to create a single case study for a project on that list. Each case will include both (1) a narrative case history to capture the richness of the case and identify any factors that determined a project's success or failure, and (2) answers to structured questions that ask about organization, technology and process issues in a consistent way across all cases.

Participants in the selected projects are being asked to complete this survey form as background information for the students to use in their projects, and we hope you can cooperate with our research.

<u>This is not a traditional questionnaire</u>. If you do not remember the details we are asking about, or if you feel that the answer would be misleading or somehow inappropriate for the project we are asking you about, feel free to leave the answer blank. You may rewrite the question so it fits better. If you have comments to add, or want to suggest a better answer than what is provided, feel free to do so.

While the students conducting this research may be cleared to discuss classified material, it should be stressed that the narratives and the answers to structured questions should never include any classified information. The results will be used in unclassified reports.

You may request a copy of any report of the findings by providing your business card, or providing a separate sheet of paper with your name and address information, including your e-mail address. If you have any questions, contact:

Richard G. Rhoades
Research Institute
The University of Alabama in Huntsville
Rhoadesr@email.uah.edu (256) 824-6343

William A. Lucas
Sloan School of Management
Massachusetts Institute of Technology
walucas@mit.edu (617) 253-0538

TO BEGIN: The first set of questions defines three dates, keyed to technology readiness levels (see page 8), and then asks about the roles played by different organizations at three stages of your project leading up to those dates. The organizations of interest are:

Prime's S&T org.:	Group within system prime contractor responsible for doing IR&D and developing new technology and concepts.
Other prime org.:	Any prime contractor organization other than the S&T organization.
Supplier S&T:	Same definition as for prime's S&T organization, but located at a supplier.
Other supplier org.:	Any supplier organization other than the S&T organization.
Army Lab/Center:	One or more of the Army laboratories or research, development and engineering Centers.
Other DoD/S&T org.:	An equivalent of an Army Lab/Center found elsewhere in DoD.

SP. What was the approximate starting date of systems planning and pre-development work? This date is **when planning work began on the integrated system**. The systems concept and applications had been formulated, but applications were still speculative. There was no proof or detailed analysis to support the approach.
SYSTEMS PLANNING START DATE (**SP**): ___/___ (mo/year) [TRL2 at system level] *See following.*

System	System planning start	Development start	Transition to production
APACHE attack helicopter	1970	1973	1982
TADS/PNVS (target acquisition and designation/pilot's night vision systems)	1976	1977	1980
MLRS rocket system	1973	1977	1980
ATACMS missile system	1980	1986	1989
M40 chemical protective mask	1982	1983	1987
Mounted microclimate cooler	1981	1982	1984
M829-A1 armor–piercing kinetic energy tank ammunition	~1983	1985	1988
TOW-2A (Tube-launched missile)	1979	1980	1984
AN/TAS 4 infrared night sight	1979	1979	1981
Joint Stars Ground Station	1973	1984	1993
Guardrail common sensor	~1980	1984	1986
PAC-2 (PATRIOT anti-missile system)	1983	1986	1990
HELLFIRE missile system	1972	1973	~1980

In what organization was the primary work leading up to this point accomplished? Prime's S&T organization, 2; Army Lab/Center, 9; Other Dod/S&T organization, 2.

Including that organization, what organizations had been involved up to this point? (Check the role of each)

Prime's S&T org.	Lead/co-lead **6**	Active support **2**	Involved **2**	Kept informed **0**	Not involved **1**	DK **2**
Other prime org.	Lead/co-lead **0**	Active support **4**	Involved **3**	Kept informed **0**	Not involved **0**	DK **6**
Supplier S&T org.	Lead/co-lead **1**	Active support **2**	Involved **3**	Kept informed **0**	Not involved **1**	DK **6**
Other supplier org.	Lead/co-lead **0**	Active support **2**	Involved **2**	Kept informed **0**	Not involved **1**	DK **8**
Army Lab/Center	Lead/co-lead **9**	Active support **3**	Involved **1**	Kept informed **0**	Not involved **0**	DK **0**
Other DoD/S&T org.	Lead/co-lead **0**	Active support **2**	Involved **4**	Kept informed **1**	Not involved **0**	DK **6**

What was the nature of the Army Lab/Center's involvement? Simulation, 4; Concept formulation, 10; Integration, 3; Requirements development, 9; Component/systems design or development, 4; Engineering support, 1.

D. Date when Development started. Typically at this date, funding started for system advanced or engineering development, a government project office was formed and prime contractor(s) selected.

DEVELOPMENT START DATE: (**D**): _____/_____ (mo/yr) *See all dates in table above.*

In what organization was the primary work in the period from SP to D accomplished? Army lab/center, 5; Supplier's S&T organization, 1; Other Prime organization, 4; Other DoD organization, 3.

What was the **Technology Readiness Level** (refer to page 15) for the SYSTEM on this date? mean = 4.69

What was the **Production Readiness** (see page 15) for the SYSTEM on this date? mean = 1.85

Including that organization, what organizations had been involved in the period SP to D? (Check the role of each.)

Prime's S&T org.	Lead/co-lead **10**	Active support **1**	Involved **0**	Kept informed **0**	Not involved **1**	DK **1**
Other prime org.	Lead/co-lead **4**	Active support **1**	Involved **3**	Kept informed **0**	Not involved **0**	DK **5**
Supplier S&T org.	Lead/co-lead **3**	Active support **5**	Involved **2**	Kept informed **0**	Not involved **0**	DK **3**
Other supplier org.	Lead/co-lead **0**	Active support **4**	Involved **4**	Kept informed **0**	Not involved **0**	DK **5**
Army Lab/Center	Lead/co-lead **6**	Active support **7**	Involved **0**	Kept informed **0**	Not involved **0**	DK **0**
Other DoD/S&T org.	Lead/co-lead **1**	Active support **4**	Involved **3**	Kept informed **0**	Not involved **0**	DK **0**

What was the nature of the Army Lab/Center's involvement? (Engineering support? Simulation or testing? Integration? Requirements interpretation?) Simulation, 8; Concept formulation, 1; Engineering support, 4; Integration, 10; requirements interpretation 1 requirements development 2 testing 1 user evaluation

TP. Date of achieving "Transition to Production" when **producible system prototype has been demonstrated in an operational environment**. Prototype is near or at planned operational system, produced on small scale. TRANSITION TO PRODUCTION (**TP**) DATE: _____/_____ (mo/yr) (TRL7 at system level) *See all dates in table above.*

What was the **Production Readiness** (see page 15) for the SYSTEM on this date? mean = 3.69

In what organization was the primary work in the period from D to TP accomplished? Army lab/center, 2; Other prime organization, 11.

Including that organization, what organizations had been involved in the period D to TP? (Check the role of each.)

Prime's S&T org.	Lead/co-lead **9**	Active support **1**	Involved **0**	Kept informed **0**	Not involved **0**	DK **3**
Other prime org.	Lead/co-lead **6**	Active support **4**	Involved **1**	Kept informed **0**	Not involved **0**	DK **2**
Supplier S&T org.	Lead/co-lead **3**	Active support **7**	Involved **1**	Kept informed **0**	Not involved **0**	DK **2**
Other supplier org.	Lead/co-lead **1**	Active support **6**	Involved **3**	Kept informed **0**	Not involved **0**	DK **3**
Army Lab/Center	Lead/co-lead **3**	Active support **8**	Involved **2**	Kept informed **0**	Not involved **0**	DK **0**
Other DoD/S&T org.	Lead/co-lead **0**	Active support **5**	Involved **2**	Kept informed **2**	Not involved **0**	DK **4**

What was the nature of the Army lab/center's involvement? Testing, 9; Simulation, 9; Integration, 3; Engineering support, 12; Requirements interpretation, 4; and Design/component development, 1.

Please note: **Here we shift away from the system as a whole, and move to its component technologies.**

T1. Now identify one or more (up to 3) technologies that were incorporated into the system you are studying. These technologies should be among those *central to the success of the system.*

System	Technology A	Technology B	Technology C
APACHE attack helicopter	Target acquisition/ designation	Computers	Missile integration (laser)
TADS/PNVS (target acquisition and designation/pilot's night vision systems)	Line of sight stabilization	Forward looking infrared target acquisition	Laser to sensor boresight alignment
MLRS rocket system	Free rocket aerodynamics	Aerodynamic pressure generating electronic fuze	Bomblet dispensing system
ATACMS missile system	Strapdown inertial guidance	Ring laser gyro	High stall torque actuation
M40 chemical protective mask	Silicone material development	Injection molding process	
Dismounted microclimate cooler **Note: Did not enter production**	Minature power source	Adaptation of compressor parts	Optimized design for tubes and vest
Mounted microclimate cooler	Y-connector design for air distribution system	Injection molding process for Y-connector	
M829-A1 armor–piercing kinetic energy tank ammunition	Penetrator design	Parasitic hardware (sabot)	Propulsion
FOG-M (fiber optic guided missile) **Note: Did not enter production**	Payout of fiber optic missile data link	TV and Infrared imaging seeker	High speed microprocessors
TOW-2A (Tube-launched missile)	Tandem warheads/ Probe development	Guidance electronics	Electro-optical countermeasures/ obscurants
AN/TAS 4 infrared night sight	Focal plane arrays (infrared)	Thermal beacon for missile	
Joint Stars Ground Station	Distributed processing	Time integration and time compression software	Raster scan monitor
Guardrail common sensor	DF location technology	Signal processing technology/ELINT	Software for sensor fusion
PAC-2 (PATRIOT anti-missile system)	Missile guidance and control	Phased-array radar	Software
HELLFIRE missile system	Semi-active laser	Spinning mass seeker head	Custom integrated circuit

T2. How new was each technology to the prime contractor? For each technology A, B, and C, was the technology:

(Answer for date you provided for Development start , D.)	**Technologies A, B and C**
New and unproven for the prime contractor?	11 of 36
Technology had been used by prime contractor but it was **new to to this kind of application**?	16 of 36
Technology had been used by prime contractor in similar applications and **was well understood**?	9 of 36
Don't know, can't remember, or would have to guess	0 of 36

T3. Production Impact. What was the impact of the technology on existing production processes?

	Technologies A, B and C
Technology forced deep and serious change?	4 of 36
Technology caused significant change?	15 of 36
Technology did not require much change	16 of 36
Don't know, can't remember, or would have to guess	1 of 36

T4. Looking back **at the Development start date,** at that time how important were these technologies to the prime?

Check (√) the best answer for each technology.

	Technologies A, B and C
This system was the Prime's only planned application of the technology.	11 of 36
Prime was planning or had started follow-on uses of the technology.	15 of 36
Technology was being used in other applications and it was expected to be significant area of competence for the Prime.	10 of 36
Don't know, can't remember, or would have to guess.	0 of 36

Now look at the SP, D, and TP dates you provided above for the System. Using the Technology Readiness Scale on page 8, find the number that represents the readiness of the separate technologies the team was working with at each point in time. Please answer here for the **state of development of each component technology**. (NOT for the over-all system which was the focus on page 1.)

	Technologies A, B and C
T5. At planning/pre-development date SP, the *technology* readiness levels were:	For N=36, average = 3.67
T6. At project development start date D, the *technology* readiness levels were:	For N=36, average = 5.25
T7. At Transition to Production date TP, the *technology* readiness levels were:	For N=36, average = 7.92

For each of the technologies A, B & C, did an Army Laboratory or Center make a significant contribution to achieving any of the above levels of technology readiness?

	Technologies A, B and C
T8. Yes, it contributed to Readiness at start of Planning/Pre-development.	22 of 36
T9. Yes, it contributed to Readiness for Development.	25 of 36
T10. Yes, it contributed to Readiness for Transition to Production.	26 of 36
Tn. No, an Army lab or center did not make a significant contribution.	4 of 36
Tdk. Don't know, can't say, don't remember.	. 0 of 36

Project History, Staffing and Location

H1. At some point, was the project either:

1. Slowed down? 5 (38.5%) 2. Stopped and restarted? 0 (0%) 3. Neither 8 (61.5%)

H2. Was the project set up as a cross-functional integrated product team (IPT), a project team drawn from different parts of the contractor's organization with most of the skills needed for the development?

1. Yes 8 (61.5%) 2. No 5 (38.5%)

If YES, was it: [for 8 cases answering yes]

1. Set up by management, with different functions & departments tasked to provide team members. 7 (87.5%)
2. Set up informally, with team expected to ask departments for help as needs emerged. 1 (12.5%)

H3. Key Skills. This question asks about "key skills" essential to the success of the project, defined as skills "that if they were not available at all, would have stopped team progress at the point when they were needed."

Were there any key skills **not adequately represented** on the team?

1. No. 8 (61.5%) 2. Yes, one. 3 (23.1%) 3. Yes, more than one. 2 (15.4%)

IF YES: H35. What were the missing key skills? Please check (✓) any and all that apply.

[for 5 cases with one or more missing skills]

Internal technical professionals	1 of 5
Producibility professionals (DFM, other)	2 of 5
Financial/contracts professionals	0 of 5
Technical/development people from Suppliers	0 of 5
Producibility professionals from Suppliers	2 of 5
Other.	3 of 5

H4. **During the Development stage** of the project, how many people on the team were collocated **very close** together? (On the same floor of a building within a one minute walk.)

1. All 2 (15.4%) 2. Most -2/3rds or more 9 (69.2%) 3. Some - over a third 1 (7.7%) 4. Few 1 (7.7%) 5. None 0 (0%)

H4a. Including the above, how many people on the team were collocated **in the same building**?

1. All 3 (23.1%) 2. Most -2/3rds or more 8 (61.5%) 3. Some - over a third 2 (15.4%) 4. Few 0 (0.0%) 5. None 0 (0%)

H5. How many people on the team involved in the **Development stage** had worked before with others on the project?

1. All 2 (15.4%) 2. Most -2/3rds or more 7 (53.8%) 3. Some - over a third 3 (23.1%) 4. Few 1 (7.7%) 5. None 0 (0%)

H6. **Whose facilities** were going to be the primary production site for the application of the new technologies?

1. Prime contractor's facilities 4 (30.8%)
2. Both prime and supplier facilities 8 (61.5%)
3. Supplier facilities 1 (7.7%)

Validation Activities: Testing and Simulation

V1. Was a failure modes and effects analysis done on the system? 1. Yes 10 (76.9%) 2. No 3 (23.1%) 3. Don't know 0

V1a. If yes, was it used to help establish the test plan? 1. Yes 9 (69.2%) 2. No 4 (30.8%) 3. Don't know 0

For **individual components**:

V2. Was there **testing to see if the individual components of the system worked**? What organization(s) did this testing? (Check for each organization that conducted this activity.)

Prime 12 (92.3%) Suppliers 12 (92.3%) Army center/lab 10 (76.9%) Other govt org. 2 (15.4%)
Not done on project 0 (0%) Don't know 0 (0%)

V3. Were there **simulations run to see if the individual components of the system worked**? What organization(s) did these simulations? (Check for each organization that conducted this activity.)

Prime 11 (84.6%) Suppliers 11 (84.6%) Army center/ lab 9 (69.2%) Other govt org. 2 (15.4%)
Not done on project 1 (7.7%) Don't know 0 (0%)

For integrated components **in controlled setting**:

V4. Were **the components tested working together in a controlled setting**? What organization(s) did this testing? (Check for each organization that conducted this activity.)

Prime 12 (92.3%) Suppliers 8 (61.5%) Army center/lab 9 (69.2%) Other govt org. 5 (38.5%)
Not done on project 0 (0%) Don't know 0 (0%)

V5. Were there **simulations of the components working together in a controlled setting**? What organization(s) did this? (Check for each organization that conducted this activity.)

Prime 10 (83.3%) Suppliers 4 (33.3%) Army center/lab 9 (75.0%) Other govt org. 2 (16.7%)
Not done on project 1 (8.3%%) Don't know 0 (0%) [N=12, one case could not be coded here]

For integrated components **in a realistic setting**:

V6. Was there **testing of the components working together in a realistic setting**? What organization(s) did this testing?

Prime 9 (69.2%) Suppliers 4 (30.8%) Army center/lab 10 (76.9%) Other govt org. 5 (38.5%)
Not done on project 1 (8.3%) Don't know 0 (0%)

V7. Was a hardware-in-the-loop type systems integration simulation laboratory used?
V7a. To see **if the individual components of the system worked:** Yes 8 (61.5%) No 4 (30.8%) DK 1 (7.7%)
V7b. To see if integrated components worked **in controlled setting:** Yes 10 (76.9%) No 2 (15.4%) DK 1 (7.7%)

V8. Recalling the total effort (100%) spent on testing and simulations, please allocate the percent of that total that were:

Average = 27% Spent to see **if the individual components of the system worked**
Average = 28% Spent to see if integrated components worked **in controlled setting**
Average = 34% Spent to see if integrated components worked **in a realistic setting**
Average = 11% Spent on any other validation purpose.
100 %

Please evaluate the following statements about the use of testing and simulations on the project.	Strongly disagree	Disagree somewhat	Neither agree nor disagree	Agree somewhat	Strongly agree	Don't know
V9. Knowledge from validation work was used consistently to improve components and system.	0 (0.0%)	0 (0.0%)	2 (15.4%)	3 (23.1%)	8 (61.5%)	0
V10. Project test philosophy was to "Break it big early."	2 (15.4%)	4 (30.8%)	3 (23.1%)	5 (23.1%)	1 (7.7%)	0
V11. Component and system maturity were validated at the right times in the program.	0 (0.0%)	0 (0.0%)	3 (23.1%)	3 (23.1%)	7 (53.8%)	0
V12. The project and the testing community had an adversarial relationship.	8 (61.5%)	3 (23.1%)	2 (15.4%)	0 (0.0%)	0 (0.0%)	0
V13. Most project validation events produced quality results.	0 (0.0%)	0 (0.0%)	4 (30.8%)	4 (30.8%)	5 (38.5%)	0
V14. The project didn't recognize important lessons that validation work uncovered.	7 (53.8%)	3 (23.1%)	2 (15.4%)	0 (0.0%)	1 (7.7%)	0
V15. Sometimes the project settled for less than the best validation method. [% of N=12 valid answers]	6 (50.0%)	5 (41.7%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	1

Team Participants & Communications **during Development**

Here are some statements about the people on the project **during the System Development stage**. Please circle a number to indicate your level of **agreement or disagreement** that each statement is a description of team processes on this project.

	Strongly disagree	Disagree somewhat	Neither agree nor disagree	Agree somewhat	Strongly agree	Don't know
D1. The team leader was good at resolving technical disagreements.	0 (0.0%)	0 (0.0%)	0 (0.0%)	7 (53.8%)	6 (46.2%)	0
D2. The team leader was good at getting necessary resources.	0 (0.0%)	0 (0.0%)	0 (0.0%)	7 (53.8%)	6 (46.2%)	0
D3. There was a lot of turn-over in team membership.	7 (53.8%)	5 (38.5%)	1 (7.7%)	0 (0.0%)	0 (0.0%)	0
D4. The team leader had both design & production experience.	0 (0.0%)	2 (15.4%)	2 (15.4%)	4 (30.8%)	4 (30.8%)	1
D5. The team leader had very high technical competence.	0 (0.0%)	0 (0.0%)	1 (7.7%)	4 (30.8%)	8 (61.5%)	0
D6. Some key technical skills were not represented on the team itself.	5 (38.5%)	4 (30.8%)	1 (7.7%)	3 (23.1%)	0 (0.0%)	0
D7. Team meetings were sometimes frustrating and non-productive.	0 (0.0%)	5 (38.5%)	4 (30.8%)	3 (23.1%)	0 (0.0%)	1
D8. Professionals were split across too many different tasks & teams.	6 (50.0%)	5 (41.7%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	0
D9. Project results did not take advantage of the team's best ideas.	6 (50.0%)	5 (41.7%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	0
D10. Key members continued through pre-production planning and testing.	6 (50.0%)	5 (41.7%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	1
D11. There was often uncertainty about the future of project funding.	6 (50.0%)	5 (41.7%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	1
D12. The team was reluctant to share concerns with government Project Manager.	6 (50.0%)	5 (41.7%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	1
D13. Management project reviews were constructive & helpful.	6 (50.0%)	5 (41.7%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	1
D14. Formal reviews were conducted at key decision points.	6 (50.0%)	5 (41.7%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	1
D15. At the prime contractor, the project was a management priority.	6 (50.0%)	5 (41.7%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	1
D16. Usually team knew right away where to get necessary outside help.	6 (50.0%)	5 (41.7%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	1
D17. Project had a visible & supportive champion in the Prime's management.	6 (50.0%)	5 (41.7%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	1
D18. There was a lot of contact with TRADOC* during the project.	6 (50.0%)	5 (41.7%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	1

* By TRADOC here and elsewhere, we mean Training & Doctrine Command and/or other appropriate user representatives.

	Strongly disagree	Disagree somewhat	Neither agree nor disagree	Agree somewhat	Strongly agree	Don't know
D19. The gov't PM was reluctant to share problems with Army leaders	6 (50.0%)	5 (41.7%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	1

D20. Who besides the team usually attended formal reviews? (Check all that apply.)

D20a. Any Prime upper management (Director or VP level)?	Yes	11	No	2	Not applicable	0	DK	0
D20b. Any Army Program management representatives?	Yes	13	No	0	Not applicable	0	DK	0
D20c. Any TRADOC or other user representatives?	Yes	12	No	1	Not applicable	0	DK	0

Activity Report during System Development Stage of Project

How often did team members do the following **during Development**? (If you feel the activity Is Not Applicable to your project, check NA.)	Never	Once or twice	Several times	Many times	Don't know Not appl
F1. Went to the shop floor to meet about related production Processes	0	0	2	10	1
F2. Asked for supplier comments & suggestions on design choices.	0	1	5	7	0
F3. Showed & discussed physical models of new components with suppliers.	0	3	4	4	2
F4. Needed management help to resolve project team disagreements.	3	7	2	1	0
How often did the following occur **during Development**?					
F5. Did TRADOC/other user organizations show strong support?	0	0	2	11	0
F6. Were there changes in key TRADOC or other user personnel?	0	7	2	3	1
F7. Were there changes in system requirements (e.g., threat)?	4	5	3	1	0

SHARED DESIGN-PRODUCTION ACTIVITIES during System Development. Here **only count joint** meetings or discussions that included **both DESIGN** and people from **PRODUCTION** and/or from the **PROGRAM** concerned with production of the System.

How often did the team members do the following **during Development**?	Never	Once or twice	Several times	Many times	Don't know Not appl
F10. Passed around physical prototypes during joint discussions.	1	0	6	5	1
F11. Held planning meetings that included both design & production people.	0	1	7	4	1
F12. Explored choices together with computational models or analytic tools.	3	2	6	1	1
F13. Had test articles or pre-production parts to discuss and examine jointly.	0	2	5	5	1

SHARED DESIGN-SUPPLIER ACTIVITIES during System Development. Now **only count joint** meetings or discussions that included personnel from **both DESIGN and SUPPLIERS.**

How often did the team members do the following **during Development**?	Never	Once or twice	Several times	Many times	Don't know Not appl
F20. Passed around physical prototypes during joint discussions.	1	2	6	2	2
F21. Held planning meetings that included both design and suppliers..	0	1	9	2	1
F22. Explored choices together with computational models or analytic tools.	2	2	8	0	1
F23. Had test articles or pre-production parts to discuss and examine jointly.	1	1	6	4	1

ACTIVITY PHASING BY STAGES OF DEVELOPMENT AND TRANSITION

WHEN were the following activities carried out by the team? For example, if in W1. Production was involved regularly in the Selection/Planning stage, dropped out, and then came back in late in the Development work and continued to participate after that, check (✓) first, fourth and fifth columns.

Please check (✓) all stages when the activity occurred.	SP Selection	D Development Early	Development Middle	Development Later	TP Transition	Never	(DK/NA)
W1. When did production representatives participate regularly? (N=12)	4	6	8	9	6	0	1
W2. When did team members meet with production on shop floor? (N=12)	0	5	8	7	6	0	1
W3. When was the TRADOC consulted on project questions? (N=13)	8	10	8	7	8	0	0
W4. When was there change in key TRADOC/user representatives? (N=13)	4	5	7	4	5	2	0
W5. When did TRADOC/other users show strong support? (N=13)	9	11	8	7	8	0	0
W6. When was there change in the system requirements? (N=13)	5	4	4	3	2	3	0

Relationship & Activities **between Engineering Design & Production/Program**

These questions are different because they focus **only** on **joint meetings or discussions** that included **both DESIGN personnel and people from PRODUCTION and/or PROGRAM** people concerned with production

	SP Selection	D Development Early	Development Middle	Development Later	TP Transition	Never	(DK/NA)
W16. When did the team & technical professionals from **Production** have unscheduled & informal **joint** conversations about the project? (N=12)	3	9	7	9	6	0	1
W17. When were analytic engineering tools used **jointly** by Design and **Production** to explore options together? (N=11)	2	2	4	7	3	3	2
W18. When were prototypes and parts used in **joint** discussions? (N=12)	0	3	8	10	7	0	1

Relationship & Activities **between Engineering Design & Suppliers**

Focus **only** on **joint meetings or discussions** that included **both DESIGN personnel** and **SUPPLIERS**:

	SP Selection	D Development Early	Development Middle	Development Later	TP Transition	Never	(DK/NA)
W26. When did the team & technical professionals from **Suppliers** have unscheduled & informal **joint** conversations about the project? (N=12)	7	9	9	8	9	0	1
W27. When were analytic engineering tools used **jointly** by Design and **Suppliers** to explore options together? (N=11)	2	6	8	6	4	1	2
W28. When were prototypes and parts used in **joint** discussions? (N=12)	0	4	9	8	7	0	1

Problem Solving and Team Effort

Here are a series of statements about problems that are said to occur with technology development. For each statement, we are asking you to make two separate judgments to help us understand what problems require substantial team effort:

❶ First, did this problem ever come up in the specific project being reported on? If "No", then circle the "0".

❷ If "Yes," how serious was the impact of this problem on the process of the project's work? Here we are concerned with **how much effort** in attention, time and energy did the project have to spend solving or compensating for this problem.

❶ Did this problem come up during this project?	No.	Yes. The problem came up. ❷ IF YES, how much project effort had to be spent on this problem? Very minor effort	Minor effort	Signif. effort	Major effort	Very major effort
B1. It was harder than expected to take the risk out of the new technology.	1	0	1	8	3	0
B2. Cut-backs in project resources forced changes/compromises.	5	3	3	2	0	0
B3. Changes in company strategies and goals hurt the project.	9	2	0	2	0	0
B4. A critical production issue was uncovered very late in the process.	5	2	3	2	1	0
B5. Management pressure pushed technology prematurely into production.	10	1	1	1	0	0
B6. There was a lack of acceptance standards for the new technology.	6	2	1	4	0	0
B7. The technology was hard to scale up from lab & pilot tests.	1	4	4	3	1	0
B8. Testing, quality control and/or acceptance took longer than planned.	3	1	2	6	1	0
B9. Departments at the prime resisted project ideas & approaches.	7	2	3	1	0	0
B10. One or more suppliers did not meet their commitments.	2	2	3	4	1	1
B11. Army Labs/Centers resisted project ideas or approaches.	8	3	2	0	0	0
B12. Army program offices resisted project ideas or approaches.	6	4	3	0	0	0
B13. Threat definition or other requirements changed during the project.	3	2	5	1	2	0

Project Outcomes

O1. Project Acceptance. Was the SYSTEM **accepted** to be put into Production? This is **initial acceptance**, not whether it actually ended up in production.

1. No, the System was abandoned.	0 (0.0%)	4. NA, not applicable	0 (0.0%)
2. No, but concept/technology was used later.	0 (0.0%)	5. DK, don't know/can't remember	0 (0.0%)
3. Yes, the System was accepted for production	13 (100.0%)		

O2. After the SYSTEM was accepted and was in Transition to Production, **how many additional changes in the designs and processes** were later required **before the System was taken into full production**?

1. Many serious changes	2 (15.4%)	4. No or almost no changes	0 (0.0%)
2. Significant changes	1 (7.7%)	5. Did not reach production, was not implemented	0 (0.0%)
3. Minor changes	10 (76.9%)	6. Not applicable 0 (0.0%) 7. Don't know	0 (0.0%)

O3. Did the SYSTEM **go into full production**?

1. No, the System was abandoned.	0 (0.0%)	8. NA, not applicable	0 (0.0%)
2. No, but concept/some technology was used later.	0 (0.0%)	9. DK, don't know/can't remember	0 (0.0%)
3. Yes, the System was put into full production.	13 (100.0%)		

O4. For each of the technologies A, B, and C above, to what extent was each used in the System as it was produced:

	Technologies A, B and C
1. No, the technology was not used in the System.	1 of 36
2. No, but the technology was used later (elsewhere)	0 of 36
3. The technology was used but not to extent originally planned.	1 of 36
4. Yes, the technology was used as planned.	34 of 36
8/9. Not applicable, Don't know.	0 of 36

O5. After the SYSTEM reached Transition to Production, **did the project go to Production as quickly as it should have**?

1. No delay	8 (61.5%)	4. Over a year late	1 (7.7%)
2. One to 6 months delay	4 (30.8%)	5. Did not reach production, was not implemented	0 (0.0%)
3. Seven to 12 month delay	0 (0.0%)	6. Not Applicable 0 (0.0%) 7. Don't know	0 (0.0%)

O6. **After the SYSTEM was actually in Production**, how many additional changes in designs and processes were required?

1. Many serious changes	0 (0.0%)	4. No or almost no changes	3 (23.1%)
2. Significant changes	2 (15.4%)	5. Did not reach production, was not implemented	0 (0.0%)
3. Minor changes	8 (61.5%)	8. Not Applicable 0 (0.0%) 7. Don't know	0 (0.0%)

O7. Did the SYSTEM as it was implemented meet the program's cost goals?

1. The results met or exceeded cost goals	6 (46.1%)	4. Did not reach production.	0 (0.0%)
2. The results came close to achieving cost goals	6 (46.1%)	8. Not applicable.	0 (0.0%)
3. The results fell far short of achieving cost goals	1 (7.8%)	9. Don't know.	0 (0.0%)

O8. Did the Development program, as implemented, come in on budget?

1. The project met or under-ran budget.	4 (30.8%)	8. Not applicable.	0 (0.0%)
2. The project slightly exceeded budget.	5 (38.4%)	9. Don't know.	0 (0.0%)
3. The project significantly exceeded budget.	4 (30.8%)		

O9. Did the System as it was implemented meet the project's technical goals and functional requirements?

1. Results met or exceeded technical goals	9 (69.2%)	4. Did not reach production.	0 (0.0%)
2. Results came close to achieving technical goals	4 (30.8%)	8. Not applicable.	0 (0.0%)
3. Results fell far short of achieving technical goals	0 (0.0%)	9. Don't know.	0 (0.0%)

O10. Did the System have problems in the field under operational conditions in Desert Storm?

1. Yes, problems in the field significantly limited the system's effectiveness.	0 (0.0%)	8. Not applicable
2. Yes, problems in the field caused minor problems in the system's effectiveness.	5 (38.5%)	0 (0.0%)
3. No, the system was deployed and encountered no noticeable loss of effectiveness.	6 (46.2%)	9. Don't know
4. No, the system was deployed and exceeded expectations of its effectiveness.	2 (15.4%)	0 (0.0%)

IF YOU CHECKED "1" or "2" to question O10, what did the field problems result from? Check all that apply.

O10a System did not meet its requirements.	1 of 5	O10d Personnel not adequately trained/prepared to use the System	3 of 5
O10b Requirements did not reflect the field environment.	4 of 5		
O10c The System was not deployed in its intended role.	1 of 5	O10e. Other reasons.	0 of 5

11. Now that you have had a chance to think about the project and provide some answers, how well do you think you feel you have captured the details of the project? Are you: (Check ✓ one.)

1. Very confident that you captured the project well? 7 (53.8%)
2. Fairly confident you understand the main things well, but not as confident about the details? 6 (46.2%)
3. Not confident of your information about the project, so we should only use your answers with caution. 0 (0.0%)

[This reflects the original judgments of the two case informants, and the student's view after he has created the master that integrates the views of the informants and other available information used in the case study.]

12. Finally, what was the most difficult problem the Project Manager faced, how was the problem dealt with, and what was the impact of the problem on the project outcome?

Case	Most Difficult Problem	Solution/Impact
APACHE	Control of production costs/influenced by integration plant location choices	Use of Army and DOD "pressure" on contractor to influence decisions/minimized impact
TOW 2A	Stability of armor threat requirements	Flexible systems engineering process that accommodated changes/minimized impact
GUARDRAIL Common Sensor	Complexity of integration of mission equipment	Use of "integrated product team" approach/minimized impact
FOG-M	Lack of sustained user support	Program could not survive development cost growth
Joint Stars Ground Station	Cost and schedule growth/delivering complex software	Additional funding and time required
TADS/PNVS	Cost growth in development	Additional funding obtained
M40 Mask	Immaturity of critical technologies	Design modified to accommodate more mature technologies
M829A1 Round	Achieving needed innovation in the system design	Design iterations employed
PAC-2	Early fielding to meet SCUD missile threat	Rapid changes in software were made
Dismounted microclimate cooler	Lack of stable user requirements due to immaturity of technology	Development program not supported
Night Sight	Selection of unqualified vendor and split management responsibility	Vendor replaced and single PMO given full responsibility
Mounted microclimate cooler	Key vendor failed to support integration schedule	RDEC used to provide expedited integration of initial units
HELLFIRE	Adversarial relationship between key vendor and prime	Army PMO staff helped to facilitate needed communications/impact minimized
ATACMS	Key vendor went out of business	Replacement vendor selected and was intensively managed by on-site senior prime contractor manager
MLRS	Establishing and managing four nation cooperative development program	Significant involvement of program leadership minimized impact

Technology Readiness Levels

1. **Basic principles observed and reported.** Scientific research begins to be translated into applied research and development concepts. There have been paper studies of technology's basic properties.
2. **Technology concept and/or application formulated.** Practical applications have been invented. Application is speculative and there is no proof or detailed analysis to support the assumptions. Examples are still limited to paper studies.
3. **Analytical & experimental critical function and/or characteristic proof of concept.** Analytical and laboratory studies have physically validated analytic predictions of separate elements of the technology. Examples include components that are not yet integrated or representative
4. **Component and/or bread board validation in lab environment.** Basic technological components are integrated to establish that pieces will work together, e.g., integration of ad hoc parts in lab. This is relatively "low fidelity" compared to the eventual system.
5. **Components and/or bread board validation in relevant environment.** Fidelity of breadboard technology is significantly increased. Basic components integrated with reasonably realistic supporting elements so the technology can be tested in a simulated environment. Examples include "high fidelity" laboratory integration of components.
6. **System/subsystem model or prototype demonstrated in a relevant environment.** Representative model or prototype system, which is well beyond the breadboard tested for TRL 5, tested in a relevant environment. Represents a major step up in a technology's demonstrated readiness. Examples include testing a prototype in a high fidelity laboratory environment or in a simulated operational environment.
7. **System prototype demonstrated in an operational environment.** Prototype near or at planned operational system. Represents a major step up from TRL 6, requiring the demonstration of an actual system prototype in an operational environment, such as in an aircraft, vehicle or space. Examples include testing the prototype in a test bed aircraft.
8. **Actual system completed and qualified in test and demonstration.** Technology proven to work in final form and under expected conditions. In almost all cases, this TRL represents the end of true system development. Examples include developmental test and evaluation in its intended weapon system to determine if it meets design specification.
9. **Actual system proven in successful operational environment.** Actual application of the technology in its final form and under mission conditions, such as those encountered in operational test and evaluation. In almost all cases, this is the end of the last "bug fixing" aspects of true system development. Examples include using the system under operational mission conditions.

Production Readiness Levels

1. The subsystem or component application embodying the technology is **produced inside the lab** by engineers, scientists or laboratory technicians to demonstrate principles for breadboard validation and testing.
2. The application is **produced outside the lab** with tools and processes used for producing very low quantities.
3. The application is produced in low quantities **using tools and processes planned to be used in production** systems. Testing procedures for components and subsystems are established.
4. The system involving the technology application(s) is **engineered for production**. All components are identified, integration, assembly and test planning is complete.
5. **Low rate production has been run** using the production processes planned for full rate production, complete with validated procedures for integration, assembly and test of the system.
6. **Full production is underway** at a scale appropriate to meet quantity requirements, testing and quality assurance has established that yields are acceptable, and the customer has accepted product from this production run.